Comparative Analysis of Cellular Proteins in Stably and Transiently Produced Lentiviral Vectors

Sabine Johnson

University College London

Submitted for the degree of

Doctor of Philosophy

I, Sabine Johnson confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Lentiviral vectors (LVs) are successfully used in clinical trials showing long term therapeutic benefits. Studying the role of cellular proteins during replication of lentivirus HIV-1 helped to understand virus assembly and budding. Knowing what cellular proteins interact with viral proteins and identifying interactions that promote formation of functional particles can be valuable for improving LV production in Gene Therapy.

The cellular protein composition of LVs produced by two different methods was compared, the transient transfection system producing vectors pseudotyped with the VSV-G envelope and a stable producer cell system producing vectors pseudotyped with the non-toxic retroviral envelope, RDpro.

Lentiviral vectors were purified using size exclusion chromatography (SEC). The number of purified LVs produced by transient transfection was six fold higher compared to stably produced particles. For linear ion trap-orbitrap tandem mass spectrometry (LTQ-MS/MS) a comparable amount of SEC purified LVs was analysed, detecting a smaller number of cellular protein species in stably compared to transiently produced vector samples.

The greater numbers of host proteins in purified transiently produced samples may due to the presence of co-purified VSV-G vesicles. On the other hand, a large number of proteins we identified had also been detected in studies of wild type viruses and HIV-1 derived vectors indicating a role in vector formation, such as viral protein transport to the assembly site in producer cells. The potential role in LV particle production of selected identified proteins was assessed. Whilst some proteins that have been detected in studies on wild type HIV-1 were found in all our samples, such as ALIX, AHNAK was unique to stably produced, RDpro pseudotyped vector samples, and thus selected for further investigation. In summary, knock down of ALIX, AHNAK and TSG101 host cell proteins in vector producer cells did not result in a significant difference in vector production.

Table of Contents

Acknowledgements

I would like to thank my supervisors Yuan Zhao and Yasu Takeuchi for giving me the opportunity to do this research project under their guidance. Their constant support and feedback throughout the last three and a half years have been vital for me to complete this work.

I also would like to thank everyone who was involved in this project at the National Institute for Biological Standards and Control, including Jun Wheeler for conducting the mass spectrometry as well as Robin Thorpe, Carl Dolman and Natalie Werling for their advice. Mary Collins and the members of her research group as well as members of the Wohl Virion Centre at the University College London have also contributed to this work by giving valuable advice in lab meetings and training in techniques, including Sean Knight, Mehdi Baratchian, Khaled Sanber and Kanayo Doi.

I express special gratitude to my parents for supporting me throughout my education. Finally, I want to thank my husband, Daniel for always being there.

Tables and Figures

Tables

Figures

Abbreviations

yc	gamma chain
2D-PAGE	two dimensional SDS-PAGE
AA	amino acid
AADC	aromatic L-amino acid decarboxylase
ABC	ammonium bicarbonate
ABCE-1	ATP-binding cassette protein family member 1
ADA	adenosine deaminase
AEXc	anion-exchange chromatography
AIDS	acquired immunodeficiency syndrome
ALD	X-linked adrenoleukodystrophy
ALIX	cell death 6-interacting protein
AP	adaptor protein
APCs	antigen presenting cells
APOBEC3G	apolipoprotein B mRNA editing enzyme-catalytic polypeptide-like 3G
AU	absorbance units
BCA	bicinchoninic acid (BCA)
BIV	bovine immunodeficiency virus
BL	Burkitt's lymphoma
BM	bone marrow
BMI1	polycomb complex protein
BSA	bovine serum albumin
CA	capsid
Ca2PO4	calcium-phosphate
CAEV	caprine arthritis encephalitis virus
cART	combination antiretroviral therapy (cART)
CAT	chloramphenicol acetyltransferase
CCND2	G1/S-specific cyclin-D2
CGD	chronic granulomatous disease
cHS4	chicken hypersensitive site-4
CID	collision induced dissociation
CIS	common integration sites
CMV	cytomegalovirus
CNS	central nervous system
COPI	coatomer complex I
cPPT	central polypurine tract
CRISRP	clustered regularly interspaced short palindromic repeats
Crm1	exportin 1
Ct	threshold cycle
CypA	cyclophilin A
DCs	dendritic cells
DEAE	diethylaminoethanol

DICER	dsRNA-specific endonuclease
Dlg1	human discs large protein
DMEM	Dulbecco's modified Eagle medium
Dox	doxycycline
DSB	double-strand breaks
dsRNA	double-stranded RNA
EBV	Epstein-Barr virus
ECL	enhanced chemiluminescence
EEF1A1 / EEF1G	eukaryotic elongation factors 1A / 1G
eGFP	enhanced GFP
EIAV	equine infectious anaemia virus
ELISA	enzyme-linked immunosorbent assay
endo-siRNA	endogenous small interfering RNAs
ER	endoplasmic reticulum
ERT	enzyme replacement therapy
ESCRT	endosomal sorting complexes required for transport
ESI	electrospray ionisation
EXP5	exportin 5
FDA	Food and Drug Administration
FIV	feline immunodeficiency virus
GALV	gibbon-ape leukaemia virus
gDNA	genomic DNA
GFP	green fluorescent protein
Ghr	hormone receptor gene
GOI	gene of interest
gRNA	genomic RNA
HBB	human ß-globin gene
HCMV	human cytomegalovirus
HDR	homology-directed repair
HEK	human embryonic kidney cells
HIV-1 or HIV-2	human immunodeficiency virus-1 or 2
HRP	horse radish peroxidase
HSCT	haematopoietic stem cell transplantation
HSCT	haemopoietic stem cells
HSPA1A	heat shock 70kDa protein 1A/1B
HSPA1A	heat shock protein
ICAM1	intercellular adhesion molecule 1
IDLVs	integration-deficient lentiviral vectors
IL2RG	IL-2 receptor common gamma-chain
imp7	importin-7
IN	integrase
IRES	internal ribosome entry site
kDa	kilo Dalton
KSHV or HHV-8	karposi's sacromas-associated virus
LAM-PCR	linear-amplification mediated PCR
LB	Luria Bertani
LC-MS/MS	liquid chromatography-linked tandem mass spectrometry
LCRs	locus control regions

LDL	low density lipoprotein
L-domain	late domain
LFA1	lymphocyte function-associated antigen 1
LTQ-MS/MS	linear ion trap-orbitrap tandem mass spectrometry
LTR	long terminal repeat
LV	lentiviral vector
m/z	mass-to charge-ratio
MA	matrix
MARCKSL1	myristoylated alanine-rich C-kinase substrate-related protein 1
MDM	monocyte derived macrophage
MDS	myelodysplasitic syndrome
mir-30	miRNA 30
miRNAs	micro RNAs
MLD	metachromatic leukodystrophy
MLMO2	*LIM domain only 2*
MLV-A	amphotropic murine leukaemia virus
MLV-RV	moloney murine leukaemia retroviral vector
MTOC	microtubule organising centre
MudPIT	multidimensional protein identification technology
MV	measles virus
MVB	multivesicular body
NADPH	nicotinamide dinucleotide phosphate
NC	nucleocapsid
NF-kB	nuclear factor kappa B
NK	natural killer
NLS	nuclear localisation signal
NPC	nuclear pore complex
Nup	nucleoporin
PBL	peripheral blood lymphocytes
PBS	phosphate buffered saline
PBS	primer binding site
PCLs	packaging cell lines
PEI	polyethylenimine
PGK	phosphoglycerate kinase
pI	isoelectric point
PIC	pre-integration complex
piRNA	PIWI interacting RNAs
PM	plasma membrane
PNS	peripheral nervous system
ppm	parts per million
PPT	polypurine tract
PR	protease
PRA1	prenylated Rab acceptor
PS	phosphatidylserine
P-TEFb	positive transcription elongation factor
Q-RT-PCR	quantitative reverse transcription polymerase chain reaction
RanGTP	small Ran GTPase
RCR	replication competent retroviruses

RER	rough endoplasmic reticulum
RISC	RNA-induced silencing complex
RNAi	RNA interference
RNAPII	RNA polymerase II
RP	reverse-phase
RPE	retinal pigment epithelium
RRE	rev-responsive element
RRV	Ross river virus
RSV	rous sarcoma virus
RT	reverse transcriptase
RT	repeat region (R)
RTC	reverse transcription complex
RVG	virus glycoprotein
SARs	scaffold attachment regions
SARs	splice acceptor
SCID	severe combined immunodeficiency
SCX	strong cation exchanger
SD	splice donor
SEC	size exclusion chromatography
shRNA	small hairpin RNA
SIV	simian immunodeficiency virus
SSFV	spleen-focus forming virus
STEM	scanning transmission electron microscopy
SU	surface
SV40	simian virus 40
T-Ag	T-large antigen
TALENs	transcription activator-like effector nucleases (TALENs)
T-ALL	T-cell acute lymphoblastic leukaemia
TAR	trans-activating response element
TBS	Tris-buffered saline
tet O	tet operator
tGFP	turbo GFP
TGN	trans-Golgi network
TM	transmembrane
Tnpo3	nuclear transport factor transportin 3
TRE	tetracycline response element
TSG101	Tumour susceptibility gene
tTA	tetracycline transactivator protein
TU/ml	transducing units per ml
UCOEs	chromatin-opening elements
VLCFAs	very long chain fatty acids
VLPs	virus like particles
VPLCs	virus producing cell lines
VSV-G	vesicular stomatitis virus-glycoprotein
VSV-G	virological synapses
WASD	Wiskott-Aldrich syndrome
WHP	woodchuck hepatitis virus
WPRE	posttranscriptional regulatory element

ZFNs	zinc-finger nucleases
ZFP	zinc-finger protein
Ψ	packaging sequence psi

1. Introduction

1.1. General Introduction

Gene therapy is defined by the American Society of Gene and Cell Therapy as a set of strategies that modify the expression of an individual's genes or that correct abnormal genes. Each strategy involves the administration of a specific DNA (or RNA). The first gene therapy trial in humans was approved by the US Food and Drug Administration (FDA) in 1990 in which a retroviral vector was used to treat two children suffering from adenosine deaminase deficiency (ADA-SCID) (Bordignon et al., 1995).

As of January 2014, over 60% of ongoing clinical trials are attempted to treat cancer, followed by monogenic disorders (9%) and cardiovascular diseases (8%). Most commonly is the use of adenoviral vectors (23.5%), retroviral vectors (19.1%) and naked plasmid DNA (17.7%). Lentiviral vectors have been used in 3.3% of all trials and are being used increasingly in newly started trials (The Journal of Gene Medicine, January, 2014).

Increasing demands for the supply of lentiviral vector require improvements in the large scale vector production, which in turn calls for a detailed understanding of vector formation in producer cells such as interactions of host proteins with viral protein and viral RNA during assembly or budding. This thesis characterises lentiviral vector associated host proteins comparing two different vector production systems. This chapter is an introduction to basic lentivirology and the development of HIV-1 based lentiviral vectors including the pioneers of retroviral vector, gammaretroviral vectors, as well as transient and stable production systems of lentiviral vectors.

1.2. Biology of Lentiviruses

1.2.1. Classification of Retroviruses

Retroviruses are a family of enveloped viral particles containing two copies of positive (+) strand RNA that is reverse transcribed and integrated into the host cell genome. According to the International Committee on Taxonomy of Viruses (Lefkowitz et al., 2013) the family of *Retroviridae* is divided into two subfamilies *Orthoretrovirinae* and *Spumaretrovirinae*. The latter consist of only one genus, Spumaviruses, including various species of Foamy Viruses, for example Simian Foamy Virus.

Orthoretrovirinae include the genera alpha-, beta-, gamma-, delta- and epsilonretroviruses as well as lentiviruses. Lentiviruses consist of human immunodeficiency virus-1 (HIV-1) and HIV-2, simian immunodeficiency virus (SIV) as well as equine infectious anaemia virus (EIAV) along with other non-primate viruses, for example feline immunodeficiency virus (FIV), bovine immunodeficiency virus (BIV) and caprine arthritis encephalitis virus (CAEV).

As the lentiviral vectors used in this study are based on the lentivirus HIV-1, this introduction will focus on HIV-1 and HIV-1 derived lentiviral vectors as well as the, to this project related, retrovirus RD114, a replication-competent feline endogenous gammaretrovirus (Reeves and O'Brien, 1984).

1.2.2. HIV-1 Genome and Life Cycle

Human immunodeficiency virus-1 (HIV-1) evolved from a simian immunodeficiency virus (SIV) and emerged in the late 20th century. It is the causative agent of acquired immunodeficiency syndrome (AIDS) (Gojobori, et al., 1990).

The viral genome of HIV-1 is a single-stranded RNA genome of approximately 9.7 kilobases and consists of nine genes that code for 15 proteins (Petropoulos, 1997). The three largest open-reading frames are common to all retroviruses and code for the major structural proteins Gag polyprotein [matrix (MA), capsid (CA), p2, nucleocapsid (NC), p1 and p6], Pol polyprotein

[protease (PR), reverse transcriptase (RT) and integrase (IN)], as well as Env [subunits: surface glycoprotein (SU) and transmembrane protein (TM)]. In contrast to gammaretroviral vectors (see section 1.3.1), LVs are complex retroviruses as their genome carries six additional genes, the regulatory genes *tat* and *rev* that are essential for virus replication and genes encoding accessory proteins Vif, Vpr, Vpu and Nef thought to modulate immune functions (Swanson and Malim, 2008). The viral genome is flanked on either side with a long terminal repeat (LTR), in particular the 5'LTR being promoter for transcription and the 3'LTR ensuring polyadenylation.

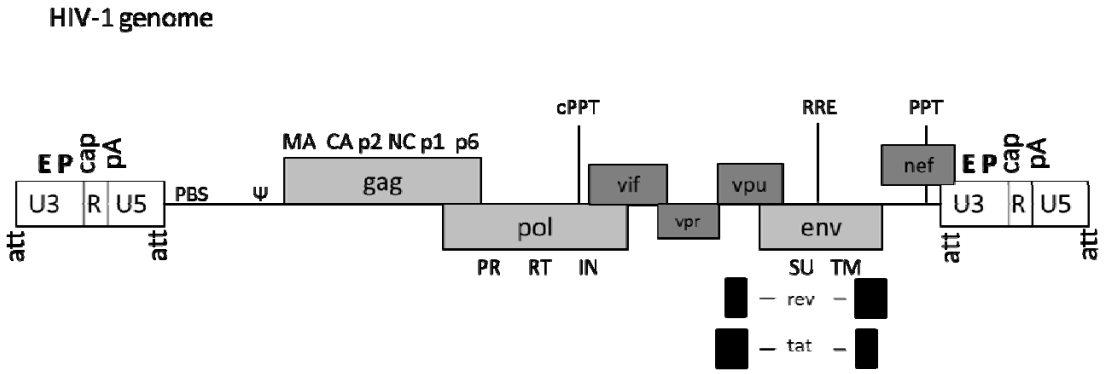


Figure 2: Integrated provirus of a HIV-1 (simplified). Adapted from (Swanson and Malim, 2008). Shown are: the long terminal repeats, 5'LTR and 3'LTR, each composed of U3, R and U5; open reading frames (light grey boxes) for Gag, Pol and Env proteins with their processed subunits, above or below, respectively. Att: integration attachment site, E: enhancer, P: promoter, cap: 5'capping site, pA: polyadenylation site, PBS: primer binding site, Ψ: packaging signal, MA: matrix, CA: capsid, NC: nucleocapsid, PR: protease, RT: reverse transcriptase, IN: integrase, SU: surface, TM: transmembrane, cPPT: central polypurine tract, RRE: Rev responsive element, PPT: polypurine tract. Regulatory genes (black boxes) rev: regulator of expression of viral proteins and tat: trans-activator of transcription. Accessory genes (dark grey boxes) vif: virion infectivity factor, vpr: viral protein R, vpu: viral protein U, Nef: negative factor. Figure is not to scale.

1.2.3. HIV-1 Virus Structure

HIV-1 virions contain the viral RNA genome, which is enclosed in a core structure surrounded by the envelope layer. Virions are on average 130 nm in diameter and contain approximately 1500 to 2500 Gag molecules determined by scanning transmission electron microscopy (STEM) of HIV-1 virions (Zhu et al., 2003, Carlson et al., 2008). The core which is a homomultimer of capsid (CA) proteins lined by matrix protein (MA) encases two single stranded RNAs, the viral genome. The RNA genome is coated by the nucleocapsid (NC) and associates with the proteins reverse transcriptase (RT) integrase (IN), protease (PR) and the viral accessory protein Vpr (Briggs and Krausslich, 2011). The envelope membrane made of cellular lipoproteins is studded with about 8 to 10 glycoproteins, each formed of three heterodimers of the surface glycoprotein (SU, gp120) and transmembrane protein (TM, gp41) (Zhu et al., 2003).

1.2.4. HIV-1 Entry

In vivo HIV-1 infects predominantly $CD4^+$ T cells and monocytes as well as macrophages. The surface glycoprotein gp120 determines this cell tropism and TM gp41 is responsible for merging of the viral and host membrane. Upon binding of gp120 to the CD4 receptor and its secondary receptor CCR5 or CXCR4, gp41 changes into its extended conformation revealing its ectodomain. The fusion peptide FP on the N-terminus of gp41 inserts into the host membrane followed by folding of gp41 into a hairpin and formation of a fusion pore with subsequent cell entry of the virus (Ashkenazi and Shai, 2011). Virions can enter the cell via clathrin-mediated endocytosis and low pH fusion with the endosomal membrane as documented by live cell imaging (Miyauchi et al., 2009).

The cellular restriction factor rhesus macaque Trim5-α has been identified to reduce HIV-1 infection of retroviruses in multiple ways, for example by acceleration of CA uncoating and thus prevention of reverse transcription (Black and Aiken, 2010). It has also been shown that Trim5-α can bind to CA of incoming retroviral cores and it was concluded that

autoubiquitinylation of Trim5-α results in subsequent degradation of the complex in the proteasome (Stremlau et al., 2004, Towers, 2007). The exact mechanisms are not completely understood and Trim5-α can possibly execute its anti-retroviral activity in several redundant ways (Malim and Bieniasz, 2012).

1.2.5. HIV-1 Reverse Transcription

After entering the cell, the reverse transcription complex (RTC) containing the viral RNA genome and proteins travels via microtubules and actin filaments to the nucleus (McDonald et al., 2002) during which the viral genomic RNA (gRNA) is converted by reverse transcription into double stranded DNA. Reverse transcriptase (RT) synthesises DNA from the viral RNA genome and duplicates the LTRs to create the full proviral DNA genome.

The HIV-1 gRNA is encompassed by the repeat region (R) on its 5' and 3' end, containing only one copy of each, U5 at the 5' and U3 at the 3' end. Minus strand DNA synthesis is initiated by binding of host cell primer $tRNA^{lysine3}$ to the primer binding site (PBS) in the gRNA (Figure 2A) (Petropoulos, 1997). Viral genomic RNA (gRNA) is copied up to the 5' end of the genomic RNA (also known as "minus strand strong stop DNA") (Figure 2B). The RNase H activity of RT cleaves the 5' end of the viral RNA and the short minus DNA strand is then transferred to the 3' end of the gRNA to continue synthesis up to the 5'PBS (Figure 2C to E). The primer for the synthesis of the plus strand is generated by cleavage of the gRNA of the RNA-DNA hybrid and is derived from the polypurine tract region (PPT). A second plus-strand DNA synthesis is initiated at the central PPT (cPPT) (Figure 2F). Initially the plus strand DNA, originating at the PPT, is synthesised up to the methylated base, 1-methyl-adenosine, of the tRNA, followed by the removal of the PPT and tRNA primer by RNase H cleavage (referred to as "plus strand strong stop DNA") (Figure 2G). This initiates the second strand transfer of the "plus strands strong stop DNA" and plus and minus strand elongation continue until both strands are copied completely (Figure 2H to I).

Apart from host cell primer $tRNA^{lysine3}$ cyclophilin A (CypA) is another host factor required for reverse transcription. CypA is a peptidyl prolyl cis/trans isomerase that catalyses the switch between *cis* and *trans* conformations of proline residue (Takahashi et al., 1989). In the context of the HIV-1 life cycle it has been shown that high amounts of CypA are packaged into virions during assembly in late stages of the replication cycle by binding to HIV-1 Gag (Franke et al., 1994). Initially it was thought that CypA incorporated during assembly could influence HIV-1 infectivity. Mutations in Gag lead to the loss of interaction of Gag with CypA in producer cells, resulting in reduced CypA incorporation but did not affect particle assembly and packaging of viral proteins or viral RNA (Braaten et al., 1996). T cells infected with CypA deficient virions contained lower levels of full length viral DNA compared to T cells infected with wild type Gag containing virions, hence a role of CypA in early stages of infection was suggested (Braaten et al., 1996). Later it was documented that CA interaction with CypA, originating in the infected cells, is more important, as knock down of CypA expression in producer cells did not affect HIV-1 virion infection (Sokolskaja and Luban, 2006). APOBEC3G (apolipoprotein B mRNA editing enzyme-catalytic polypeptide-like 3G) was shown to be another cellular restriction factor. APOBEC3G deaminates C nucleotides of the minus-strand of viral DNA during vif-negative HIV-1 replication (Harris et al., 2003). Recently it has been shown that eukaryotic elongation factors 1A and 1G (EEF1A1 and EEF1G) co-immunoprecipitate with RT and IN. SiRNA induced silencing resulted in lower reverse transcription levels as well as reduced levels of RTCs in those cells. It was suggested that EEF1A1 and EEF1G stabilise the RTC and stimulate late reverse transcription (Warren et al., 2012).

Formation of a triple DNA strand during reverse transcription, the central DNA flap, a 99 nucleotide overlap at the cPPT, (Zennou et al., 2000) has been shown to be involved in uncoating of CA from the newly formed DNA (Arhel et al., 2007). Recent data also suggests that uncoating and reverse transcription proceed in parallel potentially influencing each other (Hulme et al., 2011). The double stranded DNA complex is referred to as the pre-integration complex (PIC) and enters the nucleus via a nuclear pore.

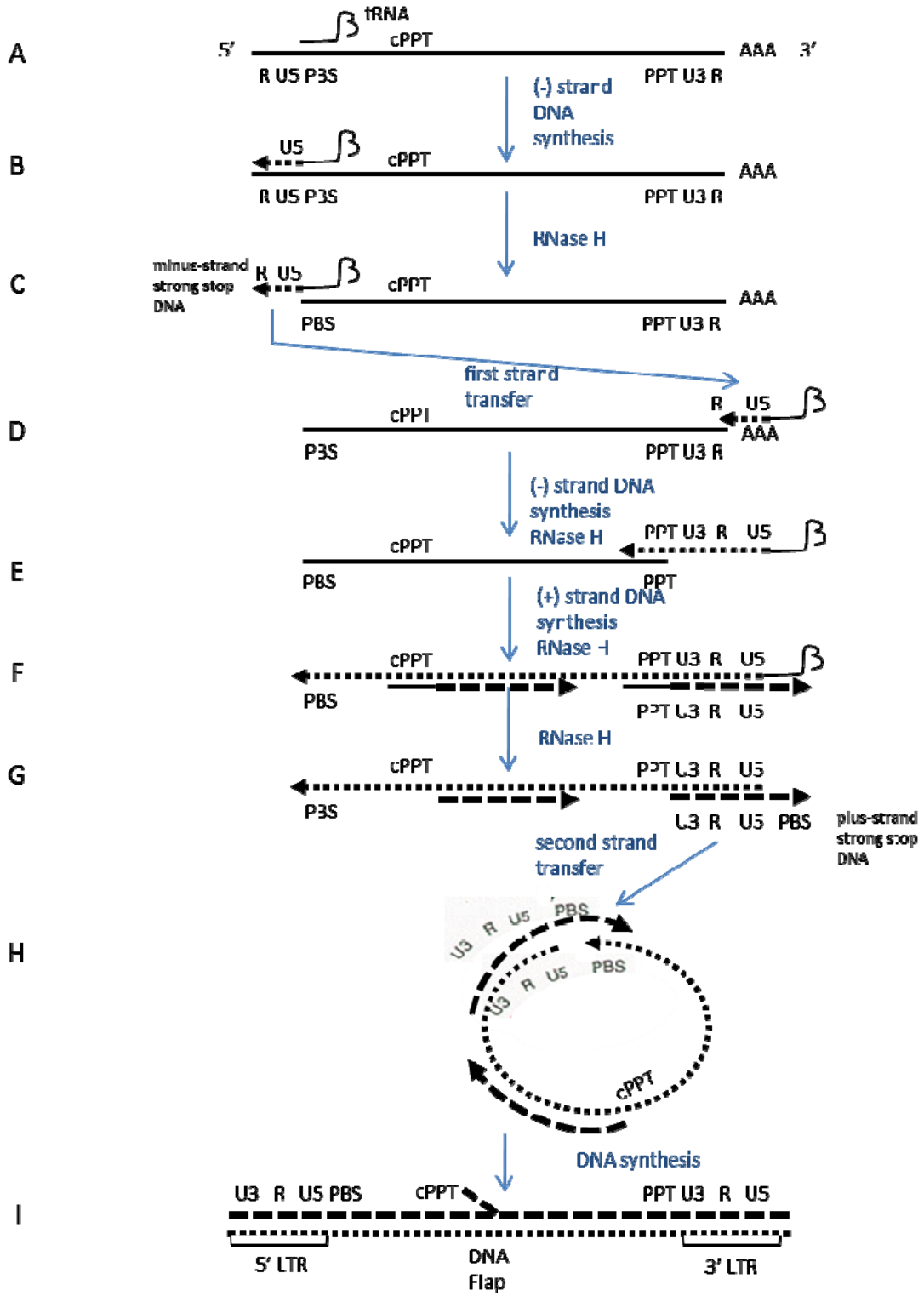


Figure 3: HIV-1 reverse transcription. Adopted from Onafuwa-Nuga et al. 2009. (A) The $tRNA^{lys3}$ is bound to complementary sequences at the primer binding site (PBS) on the HIV-1 genomic RNA (gRNA). (B) DNA minus-strand synthesis [(-)strand DNA] is initiated from primer tRNA and halts when it reaches the 5' end of gRNA, resulting in "minus-strand strong-stop

DNA." (C) RNase H degrades the template RNA strand. (D) First strand transfer: minus strand anneals to the 3'end of gRNA, complementary sequence in the R-region. (E) RNase H degrades gRNA up to PPT, minus-strand DNA synthesis continues. (F) Plus-strand DNA [(+) strand DNA] synthesis is initiated from oligoribonucleotides at PPT and cPPT. The plus-strand strong stop DNA results after synthesis stops at the first modified base in primer tRNA. (G) RNase H mediates removal of the tRNA primer. (H) Second strand transfer: plus-strand strong-stop DNA anneals at the 3' end of plus-strand strong-stop DNA to complementary sequences at the end of the minus-strand DNA. (I) DNA synthesis is completed; double-stranded DNA with long terminal repeats and a central flap at the central PPT (cPPT) is generated. The continuous line represents viral genomic RNA, dotted lines represent viral DNA. Primer tRNA is shown as beta-shaped form.

1.2.6. HIV-1 Nuclear Entry and Integration

Most gamma-retroviruses need the host cell to be in the cell cycle stage of mitosis to gain access to the nucleus through the dissolved nuclear membrane (Lewis and Emerman, 1994). Lentiviruses can infect both dividing and non-dividing cells and use the nuclear pore to enter the nucleus and access the chromatin (Suzuki and Craigie, 2007).

The PIC associates with and actively enters the nucleus through the nuclear pore complex (NPC) (Arhel et al., 2007). Generally, import of a molecule into the nucleus through the NPC involves specific transport receptors such as importin (Kutay et al., 1997). The subunit of importin-α first recognises the nuclear localisation signal (NLS) on the molecule requiring nuclear import (cargo). After binding to importin-ß, the whole complex locates to the NPC. After translocation of PIC into the nucleus, the small Ran GTPase (RanGTP) associates with importin-β inducing the release of the cargo from the receptor complex into the nucleus (Suzuki and Craigie, 2007). The exact molecular mechanisms that determine how HIV-1 crosses the nuclear membrane are still under debate; however, it has been documented that several host factors can mediate nuclear import. Depletion of the nuclear pore protein nucleoporin (Nup) 358 decreased HIV-1 infection suggesting that Nup358 helps HIV-1 to dock at the nuclear pore (Schaller et al., 2011). A genome wide screen was performed using siRNA to knock down

protein expression in target cells. Knock down of Nup358, Nup153 and the nuclear transport factor transportin 3 (Tnpo3) in HeLa cells reduced HIV-1 infectivity (König et al., 2008). Other factors identified to promote nuclear import of HIV-1 are tRNAs (Zaitseva et al., 2006) as well as importin-7 (imp7) (Fassati et al., 2003).

In the nucleus the viral DNA is integrated into the host cell DNA by the viral enzyme integrase (IN) that is associated to the viral DNA. IN binds to the attachment sites of the viral DNA LTRs, followed by an endonucleotide cleavage of the viral DNA 3'ends, called 3' processing which allows cutting of the genomic cellular DNA and simultaneous joining of the viral DNA to the 5' end of the target DNA (Krishnan and Engelman, 2012). It has been demonstrated that HIV-1 favours integration within transcribed genes (Schröder et al., 2002) and that viral gene expression levels differ significantly between different integration sites (Jordan et al., 2001).

1.2.7. HIV-1 Transcription

Viral gene expression from proviral DNA is initiated by cellular transcription factors that allow low level expression of short RNA transcripts including an RNA loop at the R-U5 region in the 5'LTR called the trans-activating response (TAR) element. These cellular transcription factors include nuclear factor kappa B (NF-kB) and transcription factor Sp1. NF-kB forms a complex with Sp1 by binding to the viral promoter in the U3 region of the 5'LTR. The short RNA transcripts are synthesised by the cellular RNA polymerase II (RNAPII) which pauses due to binding of negative elongation factors after transcription of the TAR loop and are not further elongated in the absence of Tat. Tat is a trans-activating provirus encoded protein interacting with TAR and can be expressed from a few fully elongated proviral transcripts in which RNAPII evaded pausing. As soon as sufficient amounts of Tat accumulate it can bind together with the cellular positive transcription elongation factor (P-TEFb) to TAR stem loop. This interaction allows P-TEFb to phosphorylate RNAPII, un-pausing the elongation and continuing the full RNA synthesis (Wu, 2004, Van Lint et al., 2013).

A combination of three reverse transcriptase inhibitors is the current treatment of HIV-1 infections, called combination antiretroviral therapy (cART) (World_Health_Organization, 2014). It is not fully effective in every patient and therapy interruption leads often to re-emergence of detectable viral replication. This re-emergence has been accounted to transcriptionally silent proviruses forming a latent reservoir that is thought to be caused mainly by blockage of provirus transcription post-integration. The main reservoir of latent proviruses are resting memory $CD4^+$ T cells, but macrophages and haematopoietic progenitor cells are also thought to contribute (Van Lint et al., 2013).

1.2.8. HIV-1 RNA processing, Nuclear Export and Translation

The HIV-1 genome has four splice donor sites close to the 5' end of the genome (5' splice sites) and eight splice acceptor sites (3' splice sites). More than 40 different mRNAs are produced from the pre-mRNA whilst associated with the spliceosome. These mRNAs include four incompletely spliced mRNAs encoding Vpu and Env as well as incompletely spliced mRNAs encoding Vif, Vpr and a truncated form of Tat. Completely spliced mRNAs encode Rev, Nef and full length Tat. All mRNAs include the non-coding 5'cap to the first splice donor site (5'ss D1). To facilitate nuclear export of intron containing unspliced and incompletely spliced mRNAs the HIV-1 regulatory protein Rev is required. Rev interacts with the Rev-responsive element (RRE) in the *env* gene followed by binding of additional monomers of Rev leading to oligomerisation of Rev at RRE. For nuclear export through the nuclear pore Rev interacts with karyopherin family member exportin 1 (Crm1) in the presence of GTP-bound form of the Ran GTPase. Rev is released from the mRNA and re-enters the nucleus by binding to importin-ß. After 3' processing and polyadenlyation translation of viral mRNAs is initiated. To allow translation of *gag* as well as *pol,* a -1 reading frame shift occurs at the 'slippery' sequence 200 nucleotides upstream of the *gag* termination sequence, shifting from the *gag* to the *pol* reading frame.

This occurs in 5% of translations of the unspliced mRNA resulting in one Gag-Pol precursor (Pr^{160}Gag-Pol) for every 20 Gag precursors (Karn and Stoltzfus, 2012).

1.2.9. HIV-1 Assembly

The Gag polyprotein Pr55Gag mediates HIV-1 particle assembly. It has four structural domains that all contribute to the assembly process, matrix (MA), capsid (CA), nucleocapsid (NC) and p6. MA facilitates targeting Gag to the assembly site and binding of Gag to the plasma membrane (PM). Gag dimerisation is then provided by the Gag dimerisation domain in CA followed by multimerisation of Gag through binding to viral RNA, serving as a scaffold, and followed by encapsulation of the viral genomic RNA both mediated by NC. These events and incorporation of the viral Env glycoprotein and Gag-Pol precursor protein into the viral particle can take place simultaneously along with budding of the immature viral particle from the infected cell. The latter is stimulated by p6 recruiting the endosomal sorting complex required for transport (ESCRT) (Balasubramaniam and Freed, 2011).

The identity of the HIV-1 assembly site is still under debate. Gag trafficking and the site of virus assembly, appear to be cell type dependent and are influenced by cell-cell interactions. It has been proposed that in macrophages HIV-1 particles accumulate in neutral-pH compartments (Jouve et al., 2007) connected to the PM by microchannels and are in fact invaginations of the macrophage PM (Deneka et al., 2007). Via these microchannels virions can be secreted into the cell surrounding or delivered to uninfected T cells via the virological synapse (Gousset et al., 2008). Another study suggests that the virions are assembled on late endocytic membranes (Pelchen-Matthews et al., 2003). During HIV-1 formation in tissue culture cell lines such as 293T cells GFP-tagged Gag has shown to localise at the PM (Jouvenet et al., 2006).

For targeting of Gag to the PM it binds to the highly basic region (HBR) at the surface of MA. This exposed the myristoyl moiety at the N- terminus of MA allowing it to be inserted into the lipid bilayer and stabilising the interaction between Gag and the cell membrane. The PM is rich in the anionic lipids, mainly phosphatidylserine (PS) and phosphoinositides (Saad et al., 2006)

resulting in an acidic cytoplasmid surface, allowing the basic residues in MA-HBR to interact with the acidic phospholipid PI(4,5)P2, a member of the phosphoinositides.

Viral genomic RNA associates with NC of Gag during budding to be package into particles. The NC region also catalyses dimerisation of the viral gRNA as well as annealing of the cellular tRNA primer to the PBS of viral RNA, which happens possibly during particle assembly but the exact time point is not known (Feng et al., 1999).

Unlike Gag and Gag-Pol, Env is not translated in the cytosol. It contains an endoplasmic reticulum (ER) signal sequence at its N-terminus and is synthesised on the rough ER (RER) being treated in the same way as proteins that are going to be secreted or membrane bound. After synthesis and glycosylation in the RER as a 160-kDa precursor protein (gp160) it oligomerises into trimers. Oligomers are transported to the Golgi and the trans-Golgi network (TGN), where they are cleaved by cellular proteases into mature SU glycoprotein gp120 and TM glycoprotein gp41. Noncovalently bound together they go through the secretory pathway to the PM where three molecules of each gp120 and gp41 form a heterohexameric HIV-1 glycoprotein spike and incorporate into the virus particle, with about 8 to 10 spikes per virion (Checkley et al., 2011).

The exact mechanism of how the Env glycoprotein complex is incorporated into the viral particle is under debate and currently there are four models. 1) The passive incorporation model suggesting that Env and Gag arrive independently at the PM and Env is incorporated due to its presence at the cell surface; 2) the direct Gag-Env interaction model in which gp41 of Env binds to the MA domain of Gag mediating Env incorporation; 3) the 'Gag-Env co-targeting' model suggesting a cellular structure such as the PM to which Gag and Env are both recruited; 4) the indirect Gag-Env interaction model saying that Gag and Env interact with host cell proteins that serve as linkers or adapters helping to incorporate Env (Checkley et al., 2011). As many studies have shown that cellular proteins do interact with Gag and Env during HIV-1 assembly these host proteins will be described in the next section.

1.2.9.1. Host Protein Interaction with HIV-1 during Assembly

In addition to virus encoded elements, several cellular proteins have been identified that can increase or decrease the efficiency of viral assembly. Gag and Env are recruited to the same PM site (for example a lipid raft) possibly by being connected at MA and gp41 C-terminus by a host protein. Several cellular proteins have been reported to interact with either or both Gag and Env proteins. Clathrin-associated heterotetrameric adaptor protein (AP) complexes assist normally in sorting and transporting of cellular cargo (Ohno 2006). AP-1 and AP-2 regulate the subcellular location of Env by binding to gp41 C-terminus (Berlioz-Torrent et al., 1999). AP-1u also binds Gag and silencing AP-1u by RNA interference showed a reduced Gag release in transfected cells (Camus et al., 2007). AP-3δ has been shown to help Gag trafficking to late endosomal compartments (Dong et al., 2005). The ATP-binding cassette protein family member 1 (ABCE-1) interacts with NC of Gag promoting its multimerisation at the PM (Dooher et al., 2007). TIP47, a protein involved in lipid droplet metabolism, was also shown to interact with Gag and Env and is required for their co-localisation in virus assembly (Bauby et al., 2012). Human discs large protein (Dlg1) is important in the assembly of multiprotein complexes. By binding to Gag Dlg1 modulates the subcellular distribution of Gag and HIV-1 infectivity (Perugi et al., 2009). Calmodulin can bind gp41 (Miller et al., 1993) however a clear role in HIV-1 biology has not been established yet as calmodulin also binds MA of Gag inducing the exposure of the myristoyl moiety as seen in binding Gag to the cell membrane (Tang et al., 2004).

Env gp41 also interacts with a range of other proteins. So far the following have been described: α-cantenin (Kim et al., 2002), involved in actin filament assembly; the prohibitin family members Phb1 and Phb2 (Emerson et al., 2010), associated with mitochondrial functions and cell proliferation; Luman (Blot et al., 2006), a transcription factor involved in ER stress response and prenylated Rab acceptor (PRA1), mediating trafficking of Rab proteins in vesicles (Evans et al., 2002). Further analysis is necessary to decipher implications of these interactions on HIV-1 replication.

1.2.10. HIV-1 Budding

The last phase of the virus assembly is budding from the cellular membrane involving membrane scission and particle release. The p6 domain in Gag recruits the endosomal sorting complexes required for transport (ESCRT). ESCRT serves in eukaryotes to deliver transmembrane proteins into the interior of endosomes for their eventual degradation (Raiborg and Stenmark, 2009). The ESCRT machinery consists of five heterooligomeric complexes: ESCRT-0, ESCRT-I, ESCRT-II, ESCRT-III and VPS4. The p6 domain in Gag has two binding domains for ESCRT complexes that are functionally equivalent in all retroviruses and referred to as "late assembly "or L domains. In HIV-1 the highly conserved PTAP motif within the C-terminus of p6 binds the TSG101 (Garrus et al., 2001) component of ESCRT-I and a second L domain in p6, $LYPx_nL$ binds ALIX (Strack et al., 2003) resulting in recruitment of ESCRT-III that mediates membrane scission after bud formation (Weiss and Göttlinger, 2011).

The plasma membrane contains microdomains that are enriched in sphingolipid, cholesterol and membrane anchored proteins, so called lipid rafts (Lingwood and Simons, 2010). Assembly and particle release of retroviruses including HIV-1 has been shown to take place at these microdomains (Ono, 2010). They compartmentalise cellular processes and are known to be involved in protein sorting, membrane trafficking and signal transduction (Simons and Ikonen, 1997). It was demonstrated that cholesterol depletion in HIV-1 and SIV virions harvested from T cell lines resulted in permeabilisation of virion membrane and reduced infectivity indicating that cholesterol is needed for viral membrane integrity (Graham et al., 2003).

1.2.11. HIV-1 Maturation

As viral particles are budding from the infected cell the viral protease (PR) cleaves polyproteins Gag and Gag-Pol into mature proteins. This leads to the formation of the conical core composed of the RNA dimer coated by the nucleocapsid. Nucleocapsids are encased in a capsid formed of a homomultimer of CA protein along with the enzymes RT and IN as well as Vpr (Briggs and Krausslich, 2011).

1.2.12. HIV-1 Cell-to-Cell Transmission

HIV-1 virion transmission from cell to cell has been shown to be more efficient than infection by cell-free virions (Sattentau, 2008) and can help to evade an immune response such as neutralising antibodies or complement. In HIV-1 cell-to-cell transmission, the formation of structures such as virological synapses (VS) (Pearce-Pratt et al., 1994, Jolly, 2010) and nanotubes (Sowinski et al., 2008) between retroviral producing cells and target cells have been described. VS are similar to the immunological synapse formed during antigen presentation by antigen presenting cells (APCs) to T cells and are created between APCs and $CD4^+$ T cells but also between $CD4^+$ T cells. Env glycoprotein is expressed on the infected cell and can interact with CD4 receptor and co-receptor CCR5 or CXCR4 on the target cell. Intercellular adhesion molecule 1 (ICAM1) and lymphocyte function-associated antigen 1 (LFA1) stabilise the bond. Recruitment of lipid rafts and polarisation of the secretary pathway, particularly the microtubule organising centre (MTOC) allows directing viral budding towards the target cell (Sattentau, 2008, Jolly, 2010).

1.2.13. Association of Host cell Proteins with Released HIV-1 Virions

Cellular proteins have been found to interact with viral proteins during particle formation. Some of these proteins can then be detected on the surface of the virus or within virions after they have been released from the infected cell. Studies that investigated proteins associated with HIV-1 virions can be used to identify new host cell proteins that could play a role in HIV-1 replication.

One approach to understand host cell protein and virus interactions is to purify secreted viruses and identify cellular proteins that are incorporated or on the virus surface. In density-purified HIV virions vesicles have been detected that were similar in size and composition to co-purified HIV-particles (Bess et al., 1997, Gluschankof et al., 1997). Treatment of virus preparations with subtilisin can remove almost all of these vesicles. However this procedure also removes all other surface proteins leaving only proteins inside the particles intact. As

vesicles contain the protein CD45 but not HIV-1 virions, vesicles can be depleted by treatment with anti-CD45 antibodies and only HIV-1 particles remain in the sample (Ott, 2008).

Biochemical analyses can only give total protein amounts, the distribution of proteins on individual particles cannot be measured as well as quantifying the population of virions that carries a specific protein. Studies to identify cellular proteins in HIV-1 have used also liquid chromatography-linked tandem mass spectrometry (LC-MS/MS) (Chertova et al., 2006, Saphire et al., 2006b). However the absence of a protein is not conclusive as an abundant protein can mask proteins present at a low level. Only sequences of statistical significance are considered which means that post-translational modifications, such as glycosylation that are highly variable can prevent detection by LC-MS/MS. A long list of proteins has been identified in HIV-1 particles over the last years by various techniques. The AIDS and Cancer Virus Program within the National Cancer Institute (USA) published a web-based database (http://web.ncifcrf.gov/research/avp/) containing all cellular proteins that have been identified in HIV-1 virions with the methods mentioned above and can be continuously updated by the scientific community. Cellular proteins found in secreted HIV-1 virions with a defined role in HIV-1 assembly are described below.

Certain proteins have been identified on the surface of virus particles. ICAM-1 (CD54), is known to interact with Gag shown by immunoprecipitation of the cellular protein on virus particles produced in 293T cells (Beauséjour et al., 2004). It is thought to stabilise HIV-1 binding to target cells promoting virus binding. Other proteins on the HIV-1 surface are HLA-II, Galectin-1, CD80 and CD86 as well as LFA-1 which have been shown to increase infectivity (Ott, 2008). Proteins that have been found inside HIV-1 virions include thioltransferase that removes glutathione from glutathionylated cysteine residues and can regulate protease glutathionylation of HIV-1 *in vitro* preventing its inactivation (Davis et al., 1997). The protein INI1/HSNF5 of the BAF (SWI/SNF) complex, normally responsible for remodelling of chromatin has been shown to be important for virion formation and infectivity (Sorin et al., 2006). Members of the heat shock protein (Hsp) 70 family, such as Hsp 70 and Hsp 90, have been

found in virion membranes produced by 293T cells by antibody detection and are thought to be integrated by Gag (Gurer et al., 2002)

Some proteins are known to play an active role in HIV-1 replication such as ubiquitin that is taking part in particle budding and release. Members of the actin cytoskeleton and actin-binding proteins have also been shown to be important in HIV replication. Experiments using fixed cell immunofluorescence labelling and confocal microscopy after inhibition of actin and tubulin could show reduced Gag release, infectivity and Env incorporation (Jolly et al., 2007). Another protein found to be integrated in purified virions is Staufen (Ott, 2002). Staufen is a cellular protein that is normally involved in cellular RNA localisation and was shown to interact with Gag during assembly (Chatel-Chaix et al., 2007). Staufen was also shown to bind to genomic HIV-1 RNA in virus producing cells (Mouland et al., 2000). When overexpressed in infected 293T cells its packaging into viral particles was increased as well as the amount of RNA encapsidation in each particle, however infectivity was not improved. *In vivo* Staufen potentially helps to incorporate RNA into viruses (Mouland et al., 2000). Cyclophilin A is incorporated into HIV-1 virions (Franke et al., 1994). CypA originating in the target cell is bound to CA and potentially determines the nuclear import pathway of HIV-1 post-infection (Schaller et al., 2011b). Members of the ESCRT pathway have been detected in HIV-1 virions and are known to help particle budding (see 1.2.10). Interestingly the HIV restriction factor, APOBEC3G, is actively excluded from particle packaging by the viral protein Vif (Mariani et al., 2003). 293T cells, commonly used in viral vector production (see section 1.5) do not express APOBEC3G therefore there is no need to include Vif in vector production systems (Sheehy et al., 2002).

1.2.14. Association of Host cell Proteins with Enveloped Viruses

Several cellular proteins that have been identified in HIV-1 replication or are incorporated HIV-1 virions have also been found in other enveloped viruses. Studies have been undertaken that

used mass spectrometry for protein detection. Analysed viruses include human cytomegalovirus (HCMV) (Varnum et al., 2004), Epstein-Barr virus (EBV) (Johannsen et al., 2004), karposi's sacromas-associated virus (KSHV or HHV-8) (Zhu et al., 2005) and influenza virus (Shaw et al., 2008). Members of the annexin family and heat shock protein family were detected in all of these viruses as well as cyclophilin A (not in EBV) and actin. The actin binding protein cofilin was found in HCMV, EBV and influenza virus and its expression has shown to be modified in KSHV-infected endothelial cells (McAllister et al., 2004). Another actin-binding protein moesin was detected in HCMV and EBV. These studies are further discussed in chapter 4.

1.2.15. Brief Molecular Biology of Vesicular Stomatitis Virus and RD114 Virus

In this thesis the envelope protein from VSV and the modified form of the RD114 virus envelope protein, RDpro, have been used to pseudotype lentiviral vectors (see section 1.4.5) and a brief background of their life cycles as well as known cellular proteins in VSV and RD114 is given. The envelope protein of the rhabdovirus vesicular stomatitis virus (VSV), VSV-glycoprotein (VSV-G), is commonly used in lentiviral vector pseudotyping (see section 1.4.5), as it has a high transduction efficiency and can infect a wide range of host cells (Burns et al., 1993). VSV encodes five proteins: the nucleocapsid protein (N), the phosphoprotein (P), the matrix protein (M), the glycoprotein (G), and the large polymerase protein (L). The viral core contains nucleocapsid (NC) enclosing the viral single stranded genomic RNA associated with the viral RNA polymerase (Barr et al., 2002). The low density lipoprotein (LDL) receptor has been identified as a receptor for VSV (Finkelshtein et al., 2013). The solubilised membrane phospholipid phosphatidylserine from Vero cells, kidney epithelial cells from African green monkey, was shown to saturate the binding site of VSV (Schlegel et al., 1983). Other investigators however did not see a correlation of phosphatidylserine cell surface levels and

VSV binding for various cell types (Coil and Miller, 2004). VSV enters cells by clathrin-mediated endocytosis, involving endocytic adaptor protein AP-2, Actin and Dynamin-2 (Johannsdottir et al., 2009, Cureton et al., 2009). The viral core is released into the cytoplasm after low pH-dependent fusion of the viral envelope with the endosomal membrane. The viral RNA-dependent RNA polymerase (RdRP) then transcribes copies of the viral RNA genome which are translated and new NCs are formed. NCs are transported towards the cell periphery along microtubules (Das et al., 2006). At the plasma membrane viral components assemble and bud from the cell. In a study on wild type VSV host cell proteins that are involved in virus infection were identified by genome wide RNA interference assay. Coatomer complex I (COPI), a protein complex involved in vesicular transport from the Golgi apparatus to the endoplasmic reticulum (ER) was identified and it was hypothesised that depletion of COPI negatively influences VSV-G protein transport to the PM (Panda et al., 2011). Hence expression levels of this protein could influence vector production of VSV-G pseudotyped viral vectors. Cyclophilin A was shown to be required for VSV life cycle; overexpression of a mutant form of Cyclophilin A in infected cells resulted in reduced virus replication (Bose et al., 2003).

RD114 is a replication-competent feline endogenous gamma-retrovirus. The receptor for RD-114-type viruses in humans has been identified as sodium-dependent neutral amino acid transporter-type 2 (Tailor et al., 1999) and is also highly expressed on $CD34^{+}CD38^{-}$ cells (Brenner et al., 2003) making RD114 glycoprotein suitable for pseudotyping of viral vectors targeting multipotential haemopoietic stem cells (HSC). It has been suggested that RD114-Gag recruits the ESCRT machinery to the PM for viral assembly similar to other retroviruses including HIV-1 (Votteler and Sundquist, 2013). RD114 budding was greatly inhibited by alanine substitution of the PPPY L-domain in pp15 of RD114-Gag (Fukuma et al., 2011). Literature search indicates that RD114 virion-associated host proteins have not been studied to date.

1.3. Gammaretroviral Vectors

1.3.1. Gammaretroviral Genome and Particle Structure

The integrated gammaretroviral provirus is flanked by two LTRs, composed of U3, R and U5. Viral transcription is initiated from the 5'U3 region containing the enhancer and promoter. Thus the viral RNA starts with the 5'R region, followed by U5, the primer binding site (PBS), the major splice donor (SD), the packaging and RNA dimerisation signal (Ψ). The latter is located upstream of viral protein encoding genes *gag, pol* and *env* and splice acceptor (SA) in *pol*. Downstream of *env* are the polypurine tract (PPT) and the 3'U3 and 3'R region. The 3'R region contains the polyadenylation site which gives the viral RNA a poly A tail resembling a cellular mRNA (Figure 3).

Unspliced viral RNA will be packaged into particles along with viral proteins encoded by the open reading frames. These included the structural proteins matrix (MA), p12, capsid (CA) and nucleocapsid (NC) encoded by *gag*, proteins for virus replication, protease (PR), reverse transcriptase (RT) and integrase (IN) encoded by *pol* and the envelope protein. Proteins encoded by *gag* and *pol* are cleaved during virion maturation by the viral protease whereas Env is cleaved into surface (SU) and transmembrane (TM) by a cellular protease during transport to the assembly site in the Golgi apparatus (Maetzig et al., 2011). Compared to the complex lentiviral RNA genome, gammretroviruses have a simple RNA genome only encoding for Gag-Pol and Env and lacking genes encoding for accessory and regulatory proteins. Only one spliced mRNA encoding Env and no multiple spliced mRNAs are produced (Coffin et al., 1997).

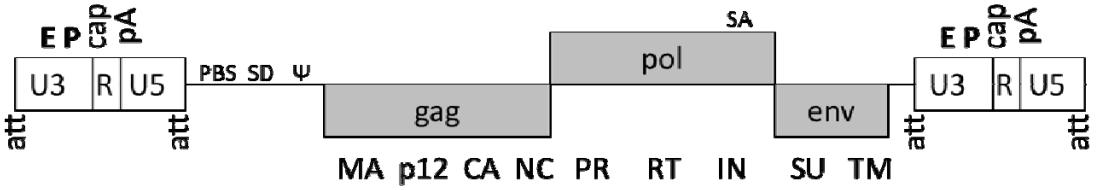


Figure 4: Integrated provirus of a gammaretrovirus (example MLV, simplified). Adapted from (Baum et al., 2006). Depicted are the long terminal repeats 5'LTR and 3'LTR composed of U3, R and U5 and open reading frames (filled boxes) for Gag, Pol and Env proteins with their processed subunits shown below. Att: integration attachment site, E: enhancer, P: promoter, cap: 5'capping site pA: polyadenylation site, PBS: primer binding site, SD: splice donor, Ψ: packaging signal, SA: splice acceptor, PPT: polypurine tract, MA: matrix, CA: capsid, NC: nucleocapsid, PR: protease, RT: reverse transcriptase, IN: integrase, SU: surface, TM: transmembrane. Figure is not to scale.

1.3.2. Development of Gammaretroviral Vectors

1.3.2.1. Split Packaging Design

Retroviruses insert their viral genome into cells they infect followed by expression of the viral proteins, hence they can be exploited as a delivery vehicle for therapeutic genes to target cells. The therapeutic gene is also referred to as the gene of interest (GOI) and is inserted into the viral genome. Research on retroviruses for their use as gene therapy vectors started in the early 1980s (Mann et al., 1983) and since then several modifications to the original viral genome have been made. The development of retroviral vectors as a gene delivery system has been especially focused on increasing their safety for their use in clinical trials. To avoid generation of replication competent retroviruses (RCR) by homologous recombination or template switching during reverse transcription the split packaging design was developed. The viral genome is divided and expressed from three separate DNA plasmids. Two plasmids express the viral genes *gag-pol* and *env*, respectively, from a promoter. The third plasmid, the

transfer vector, expresses the therapeutic gene flanked by LTRs harbouring the PBS. The *Gag-Pol* backbone is devoid of the packaging signal Ψ and vector coding sequences and only the vector genome RNA that contains the packaging signal Ψ is packaged into vector particles.

1.3.2.2. Further Improvements in Vector Safety and Efficiency

The conventional transfer vector contains a transgene cassette flanked by two LTRs. Expression from RNA transcripts in producer cells starts and ends with R regions as transcription termination is ensured by a polyadenylation in the 3'R region. After transduction of target cells with the produced vectors the 3'U3 region is copied to the 5'LTR during reverse transcription (section 1.2.5).

Several potential scenarios using this vector design raised safety issues. If transduced cells are infected with wild-type retroviruses, after transduction with a retroviral vector with full length LTRs, the HIV-1 wild-type virus could act as a helper virus and package the integrated vector into new viral particles spreading them beyond the target cells. Another issue can be the activation of cellular genes including oncogenes from the strong enhancer and promoter in 5'U3 region. These concerns led to the first development of an MLV based self-inactivating (SIN) vector containing a deletion of 299 base pairs in the 3'U3 region, including the enhancer and promoter sequences (Yu et al., 1986). Later on problems with this design were identified. The deletion of this large part of the 3'U3 region increased the probability of read-through of transcription in MLV-SIN and HIV-1 -SIN vectors beyond the R' region, suggesting that U3 contains termination enhancer motifs in addition to enhancer and promoter sequences (Zaiss et al., 2002). To circumvent this safety risk the heterologous polyadenylation enhancer element derived from simian virus 40 (SV40) was introduced in the 3' deletion to act as a termination enhancer (Schambach et al., 2007). Furthermore SIN-gammaretroviral vectors were of lower titers compared to those with full length LTRs (Schambach et al., 2006)

In order to find the optimal gene expression level and produce high titer vectors several other modifications have been made. Early γ-RVs expressed the transgene from the 5'LTR, harbouring enhancer and promoter sequences in the U3 region. In improved variants the gammaretroviral U3 region was replaced with those of spleen-focus forming virus (SFFV) showing higher transgene expression (Baum et al., 1995). The introduction of a minimal splice acceptor oligonucleotide into the HIV leader sequence creates an artificial intron, downstream of the 5'LTR and upstream of the transgene, which was also shown to increase GOI expression (Hildinger et al., 1999). Retroviral vectors can be combined with envelope proteins of other viruses, so called pseudotyping, mediating target cell specificity. A more detailed description is outlined in the context of lentiviral vector pseudotyping in section 1.4.5.

1.3.3. Gammaretroviral Vectors in Gene Therapy Trials

1.3.3.1. Severe Combined Immunodeficiency (SCID)

1.3.3.1.1. ADA-SCID

ADA-SCID is caused by mutations in the enzyme adenosine deaminase (ADA). Ubiquitously expressed in all tissues of the body, ADA converts the toxic metabolites deoxyadenosine and adenosine from broken down DNA and RNA, respectively. ADA deficiency leads to the accumulation of these toxic metabolites resulting in impaired development of functional T, B and natural killer (NK) cells. The symptoms of the resulting severe combined immunodeficiency (SCID) are repeated and persistent infections from early childhood onwards and can ultimately lead to death when left untreated. Besides allogenic haematopoietic stem cell transplantation (HSCT), enzyme replacement therapy (ERT) in the form of PEG-ADA intramuscular injections is the treatment of choice. Alternatively gene therapy treatment was first tested in the early 1990s.

In early trials a γ-RV was used encoding an ADA-minigene in the 3' LTR to result in two expression cassettes per integrated vector after duplication of the 3'LTR U3 during reverse transcription (Hantzopoulos et al., 1989). Autologous bone marrow (BM) cells and peripheral blood lymphocytes (PBL) were transduced and injected intravenously. One year after the treatment PBL-derived cells containing the vector declined and were progressively replaced by BM derived gene-marked cells (Bordignon et al., 1995). In this patient, PEG-ADA treatment was continued so that clinical benefits due to gene therapy treatment were difficult to determine. Another trial showed that PEG-ADA discontinuation led to preferential expansion of gene-marked T cells containing the ADA gene (Bordignon et al., 1995).

In a later trial two patients, for whom PEG-ADA was financially not available, were treated with the γ-RV GIADAl encoding ADA cDNA under the control of the LTR. Transduction of autologous HSC after mild pre-conditioning resulted in reconstituted T cell proliferative response and normalised serum immunoglobulin levels in both patients with no respiratory infection up to 12 months after treatment (Aiuti et al., 2002). Eight additional patients were treated. Follow up studies showed that a high percentage of transduced T, B and NK cells between 54% and 88% showed successful reconstitution of the immune system one year after treatment. After a median follow up of four years, eight out of a total of ten treated patients did not need PEG-ADA therapy any longer. A polyclonal T cell receptor repertoire implied that no clonal dominance had occurred that is normally caused by vector insertions near cellular proto-oncogenes (Aiuti et al., 2009). Five more patients have been treated with this protocol and all fifteen patients were well with only two patients requiring ERT after gene therapy (Ferrua et al., 2010).

In a trial in the UK six patients were treated with nonmyeloablative conditioning and γ-RVs delivering human ADA cDNA driven by the SFFV LTR. These vectors also contained woodchuck hepatitis virus (WHP) posttranscriptional regulatory element (WPRE) for increased transgene expression. After a median follow up of 43 months, metabolic detoxification in three patients with sustained ADA expression from different haematopoietic lineages and functional gene-

marked T cell levels were observed. Two patients had to restart ERT permanently due to graft failure. Similar insertions into the MDS1 and EVI1 complex (MDS-EVI1) locus, encoding a transcriptional regulator and oncoprotein, like in the CGD trial (see section 1.3.3.2) were seen (Stein et al., 2010). So far adverse effects such as the development of leukaemia due to vector mediated insertional mutagenesis have not been seen in any of the patients (Gaspar et al., 2006, Gaspar et al., 2011).

To date all of the ADA-SCID gene therapy patients worldwide survived the treatment and an overall efficacy shown by cessation of ERT in more than 70% of patients (Gaspar, 2012) has been documented. Nevertheless, all the conducted trials used γ-RVs with intact viral promoters that are known to have caused insertional mutagenesis in other gene therapy trials (Stein et al., 2010, Hacein-Bey-Abina et al., 2008) and therefore treated individuals will need to be closely monitored in the future. To reduce the possibility of γ-RVs insertional mutagenesis SIN lentiviral vectors are tested as an alternative. A multicentre trial is ongoing using a SIN-LV under the control of the short form of the elongation factor1α promoter (EFS) (Mukherjee and Thrasher, 2013).

1.3.3.1.2. X-linked SCID

X-linked SCID accounts for 50% to 60% of all reported SCID cases and is caused by mutations in the common gamma chain (yc) gene a subunit shared by several cytokine receptors (IL2, 4, 7, 9, 15, and 21R, located on the X-chromosome. Patients are unable to develop T and NK cells and have defective B cells (Fischer, 2000) resulting in recurrent infection of the respiratory and gastrointestinal tracts and the failure to thrive (Buckley et al., 1997). Bone marrow transplantation is the only available treatment providing a partially or fully matched donor is available.

The first clinical trial for X-linked SCID in Paris included 10 patients treated with autologous bone marrow derived CD34+ cells transduced with a moloney murine leukaemia retroviral

vector (MLV-RV) delivering yc cDNA driven by the MLV LTRs (MFGyc vector) (Hacein-Bey et al., 1996, Cavazzana-Calvo et al., 2000, Hacein-Bey-Abina et al., 2002). No adverse events were observed up to one year after gene therapy. T cell levels were normal in six patients, however four patients developed T-cell acute lymphoblastic leukaemia (T-ALL), of which one patient died. Insertional analysis showed that cells of those four patients harboured vectors at predominant integration sites near the transcription start site of the proto-oncogenes *LIM domain only 2* (*LMO2*), *G1/S-specific cyclin-D2* (*CCND2*) *or Polycomb complex protein* (*BMI1*) (Hacein-Bey-Abina et al., 2008). The latest follow up study 8 to 11 years after treatment on eight of the patients shows that all are well including one individual that had received HSCT (Hacein-Bey-Abina et al., 2010).

A second trial was conducted in London with ten patients receiving autologous bone marrow derived CD34+ cells transduced with the same vector used in the Paris trial but pseudotyped with gibbon-ape leukaemia virus (GALV) envelope instead of amphotropic MLV. Twelve to 29 months after treatment of the first four treated patients all had normal T cell counts, detectable but low levels of NK cells and no adverse effects were observed (Gaspar et al., 2004). Within two years of gene therapy one out of a total of ten treated patients developed T cell acute lymphoblastic leukaemia (T-ALL). Peripheral blood cell analysis identified a single insertion near the transcription start of the proto-oncogene *LMO2* resulting in its overexpression as well as that of neighbouring genes. Expression changes of other genes were also observed, such as down-regulation of tumour suppressors *p14(ARF1)* and *p16(INK4a)* and overexpression of leukaemia associated *NOTCH1* suggesting that a combination of genetic changes additionally to insertional mutagenesis contributes to the development of leukaemia in this patient (Howe et al., 2008). The same vectors (MFGyc pseudotyped with GALV or amphotrophic MLV) were used in separate trials in London and Paris as well as the USA in five older patients (10 to 20 years) but did not show any clinical benefits (Thrasher et al., 2005, Chinen et al., 2007).

In summary, X-linked SCID can be successfully treated by gene therapy however the development of leukaemia required further vector development and a shift to using lentiviral vectors (see section 1.4.1).

1.3.3.2. Chronic granulomatous disease (CGD)

Chronic granulomatous disease (CGD) is a rare primary immunodeficiency characterised by severe and life-threatening bacterial and fungal infections and tissue granuloma formation caused by deficient antimicrobial activity of phagocytes. This defect is caused by mutations in any of the five genes encoding phagocytic oxidase forming the nicotinamide dinucleotide phosphate (NADPH) oxidase complex.

In the USA a total of thirteen patients with CGD was treated with γ-RVs encoding phagocytic oxidase with or without pre-conditioning however only 0.13% of CD34+ cells were gene-corrected and enzyme activity declined after 8 months resulting in no long-term clinical benefits (Malech et al., 1997, Malech, 2000, Kang et al., 2010, Mukherjee and Thrasher, 2013).

In another two trials in Germany and Switzerland, pre-conditioning with busulfan was used before gene therapy in two patients per trial with X-linked CGD (Ott et al., 2006, Stein et al., 2010, Bianchi et al., 2009, Bianchi et al., 2011). In one trial up to 50% of total neutrophils in patients peripheral blood contained the integrated γ-RV encoding the *gp91^{phox}* gene driven by spleen focus forming virus (SFFV) derived LTR. This was found to be caused by clonal expansion of cells with vectors inserted into oncogenic loci MDS-EVI1, PRDM-16 and SETBP1 (Ott et al., 2006) and both patients developed myelodysplasitic syndrome (MDS), characterised by the ineffective production (or dysplasia) of the myeloid cell lineage. Furthermore methylation of the retroviral promoter caused loss of oxidase function (Stein et al., 2010). In the second trial one patient developed MDS and the second patient showed clonal expansion of transduced neutrophils (Mukherjee and Thrasher, 2013).

In patients treated in London less than 10% of neutrophils carrying the transduced gene were detected and became undetectable by day 42 post treatment (Mukherjee and Thrasher, 2013). Two patients treated in Seoul had only 0.2 to 0.9 % of gene corrected cells in peripheral blood one year after treatment (Kang et al., 2011). A clinical benefit after treatment in some of the patients is encouraging for the use of gene therapy in CGD patients however apart from improvements in the level of gene-marked cells, improvements in vector design need to be undertaken to avoid integration into sites near to or within transcription starts of proto-oncogenes. As an alternative to γ-RV lentiviral vectors are being tested in a currently ongoing trial of CGD sponsored by Genethon, for the treatment of 20 patients over 4 different sites, including the UK, Germany, Switzerland and France (Genethon, 2013, ClinicalTrials.gov, 2013).

1.3.3.3. Wiskott Aldrich Syndrome (WAS)

Wiskott-Aldrich syndrome (WAS) is a primary immunodeficiency with mutations in the *WAS* gene. *WAS* regulates the cytoskeleton and WAS-deficiency deregulates proliferation and activation of immune cells leading to recurrent infections, autoimmunity, thrombocytopenia (low platelet count) and lymphoid malignancies. Ten patients were treated using a GALV-pseudotyped γ-RV with full length LTRs following preconditioning. An initial report showed restored gene expression and a proportion of corrected lymphocytes as well as reduced frequency and severity of infections (Boztug et al., 2010) however seven patients developed leukaemia and insertional integration and transactivation of proto-oncogenes MDS1-EVI1, PRDM16, LMO2 and CCND2 was observed (Braun et al., 2014).

1.4. Lentiviral Vectors

1.4.1. Advantages of Lentiviral over Gammaretroviral Vectors

In recent clinical trials LVs are being used preferentially over γ-RVs as they have several advantages. Integration profile analysis of patient's samples treated with γ-RVs showed that they tend to integrate close to the transcription start site and near or into proto-oncogenes as shown by analysis of integration sites by linear-amplification mediated PCR (LAM-PCR), sequencing unknown DNA flanking sequences, in CGD patients (Ott et al., 2006). It is generally accepted that the enhancers and promoter sequences in 5'U3 region of the integrated provirus in target cells in these vectors transactivate these oncogenes leading to clonal expansion that resulted in leukaemia in some patients. Integration site analysis results of the latest clinical trial using LVs in patients with Wiscott-Aldrich syndrome (WAS) showed that these vectors cluster in gene-rich regions and integrate within transcriptional units but integration sites covered a wide variety of biological categories and did not dominate proto-oncogenes (Aiuti et al., 2013). Follow up of patients showed similar integration site profiles compared to results of an earlier X-linked adrenoleukodystrophy (ALD)-trial using LVs (Cartier et al., 2009) without clonal expansion or leukaemia. Furthermore γ-RVs can only infect cells that are going through mitosis, hence are proliferating (Lewis and Emerman, 1994). When using γ-RVs HSCs need to be induced using a cocktail of cytokines and other growth factors to enter cell cycling prior to transduction (Cornetta et al., 2008). In contrast LVs can transduce non-dividing cells (Naldini et al., 1996) by entering through the nuclear pore. Additionally LVs have a larger transgene capacity and are suitable to deliver large therapeutic genes. Insert with a size of up to 14 kb have been tested however titers reduce with increased insert size (Kumar et al., 2001).

1.4.2. Lentiviral Vectors in Gene Therapy Trials

1.4.2.1. Acquired Immunodeficiency Syndrome (AIDS)

Lentiviral vectors have been used the first time in a clinical trial for the treatment of Acquired Immunodeficiency Syndrome (AIDS) starting in 2003. Autologous $CD4^+$ T cells from AIDS-patients were transduced with a lentiviral vector, named VRX496, carrying an antisense sequence of HIV *env* (Lu et al., 2004) in order to decrease viral load and increase the number of $CD4^+$ T cells (MacGregor, 2001). In a follow up study, one year after treatment, transduced cells could engraft in all five patients resulting in an increased number of $CD4^+$ T cells in four patients and a decrease in viral load in one out of five patients (Levine et al., 2006). By 2013 another 60 patients have been treated with VRX496. Clinical follow up results of eight patients of whom records of viral loads before treatment with combination antiretroviral therapy (cART) were available were published. Following infusion of VRX496 transduced autologous T cells patients were taken off cART and as expected the viral load increased, but stayed below levels before cART treatment in six out of eight patients. A high variability in persistence of gene-modified PBMCs was seen with a maximum of 10% in two patients (Tebas et al., 2013). For a more informative outcome follow up results of all patients need to be seen.

1.4.2.2. X-linked Adrenoleukodystrophy (ALD)

Lentiviral vectors were also used for the treatment of X-linked adrenoleukodystrophy (ALD), a severe brain demyelinating disease. Mutations in the *ABCD1* gene lead to a deficiency in ALD protein, an adenosine triphosphate–binding cassette transporter that helps to degrade very long chain fatty acids (VLCFAs) in oligodendrocytes and microglia. In patients with mutations in the *ABCD1* gene VLCFAs accumulate in the peroxisomes of these cells disrupting myelin maintenance. In a clinical trial two patient's autologous CD34+ haematopoietic progenitor cells were transduced with lentiviral vectors delivering the wild type *ABCD1* cDNA expressed from the MND promoter (myeloproliferative sarcoma virus enhancer) (Cartier et al., 2009). Twenty months post-transplantation of the transduced CD34+ cells VCLFA levels were reduced 20%

and 28% in patient 1 and 2 respectively. Of all peripheral blood mononuclear cells (PBMCs) 10% and 15% still expressed vector delivered ALD after 30 months of treatment. Integration site analysis by linear amplification mediated (LAM)–PCR showed insertions into the same gene or genomic region in two or more individual cell clones, termed common integration sites (CIS) however no haematopoietic clones dominated.

1.4.2.3. Beta-Thalassaemia

Another clinical trial with clear therapeutic benefits after using lentiviral vector gene therapy has been undertaken in two patients with beta-thalassaemia. Beta-thalassaemia is a group of blood disorders caused by mutation of the *HBB* gene encoding human ß-globin. These mutations can be hetero- or homozygous resulting in reduced or absent levels of ß-globin polypeptide chains in red blood cells leaving them unable to bind α-globin chains to ß-globin chains to form haemoglobin A. The current treatment options for beta-thalassaemia are regular blood transfusions or HSC transplantation. In the gene therapy trial patients were transplanted with autologous CD34+ cells carrying a ß-globin gene variant, $\beta^{A(T87Q)}$ which had been delivered by lentiviral vector transduction. Beta-globin protein derived from the gene variant will be distinguishable in the patients' blood during treatment follow up studies. One patient failed to engraft the transplanted CD34+ cells due to the technical handling of the cells (Cavazzana-Calvo et al., 2010). The follow up study of the second patient up to three years after treatment showed that nucleated blood cells increased to 11% and left him transfusion independent with stable levels of haemoglobin one year post-treatment. Clonal expansion of cells with vector integrated into the cell genome at the high mobility group protein locus, *HMGA2* had been observed without evidence of leukaemia development (Cavazzana-Calvo et al., 2010). However, six years after the trial this clone is not dominant anymore and ß-globin expression is still sufficient for the patient to remain transfusion independent (Leboulch, 2013).

1.4.2.4. Metachromatic Leukodystrophy (MLD)

Metachromatic leukodystrophy (MLD) is an early onset neurodegenerative lysosomal storage disorder caused by mutations in the *ARSA* gene, resulting in a deficiency of the enzyme arylsulfatase. The enzyme substrate sulfatide accumulates in cells of the central nervous system (CNS) and Schwann cells and macrophages of the peripheral nervous system (PNS). In 2010, a clinical trial phase I/II was undertaken treating 10 children with *ARSA* deficiency with gene therapy (Verma, 2013). The results of three patients are reported. A third generation lentiviral vector (LV) (see section 1.4.3) delivering the correct form of the human *ARSA* cDNA under the control of phosphoglycerate kinase (PGK) promoter (Biffi et al., 2013) was used. After pre-conditioning treatment using myeloablative bulsufan autologous bone marrow derived CD34+ cells transduced with the LVs were re-administered. A high-level of stable engraftment was achieved with 40% to 80% of bone-marrow derived haematopoietic colonies containing the vector. Common insertion sites were found to be the same as in the ALD HSC-GT trial (Cartier et al., 2009) confirming a preference for integration within transcriptional units. However, no clonal dominance was seen in these patients. Gene expression of functional ARSA was reconstituted to above normal levels. Follow up from 18 to 24 months showed that two patients remained asymptomatic and the third patient's pre-existing severe peripheral neuropathy improved after the therapy (Biffi et al., 2013).

1.4.2.5. Wiskott Aldrich Syndrome (WAS)

As described in section 1.3.3.3, Wiskott-Aldrich syndrome (WAS) is a primary immunodeficiency with mutations in the *WAS* gene. Coordinated by the same institute that undertook the MLD trials, gene therapy using lentiviral vectors was applied in six patients with WAS, the results of three patients are reported. After pre-conditioning to deplete endogenous HSCs, LV-transduced autologous CD34+ cells were re-administered. LVs were self-inactivating (see section 1.3.2.2) delivering the correct form of human WAS under the control of the WAS promoter (Aiuti et al., 2013). Engraftment levels of 25% to 50% in HSC in the bone marrow

were achieved. WAS protein expression was restored to normal levels as well as immunological responses such as T cell proliferation and NK cell cytotoxic activity. Analysis of IS showed a preference for transcriptional units and gene dense regions and similar CIS in comparison to the ALD trial. No clonal dominance was detected 6 to 18 month after treatment. Pre-treatment eczema dissolved after 6 to 12 months of treatment and frequent infections decreased.

1.4.2.6. Acute Lymphoid Leukaemia

Two children with acute lymphoblastic leukaemia (ALL) were treated with T cells transduced with LVs delivering the chimeric antigen receptor CTL019 including a CD137 (4-1BB) signalling domain. Transduced T cells expressed a modified receptor able to kill CD19 expressing tumour cells. Complete remission of ALL and robust expansion of CTL019 expressing T cells was observed in both patients without clonal expansion of a dominant T cell receptor clone and was ongoing in one patient up to 11 month follow up. The third patient had a relapse of leukaemia with CD19 negative tumour cells showing that in some patients more than one target molecule needs to be delivered to decrease the likelihood of emergence of CD19 negative escape cells (Grupp et al., 2013).

1.4.2.7. Parkinson's Disease

Recently the result of a first clinical trial using a lentiviral vectors for the treatment of 15 patients with Parkinson's disease have been published (Palfi et al., 2014). A VSV-G pseudotyped tricistronic EIAV-based vector, named ProSavin (Oxford BioMedica, Oxford, UK), was administered *in vivo* into the striatum of the brain targeting postmitotic neurons. Vectors delivered the rate-limiting dopamine biosynthetic enzymes tyrosine hydroxylase, aromatic L-amino acid decarboxylase (AADC) and cyclohydrolase driven from CMV and two IRES, respectively. Sustained transgene expression was reported. Follow up studies at 12 months showed mild drug-related adverse events, with the majority being increased dyskinesias (involuntary muscle movements), which resolved after reduction of the oral dopaminergic

medication suggesting an efficacious dopaminergic therapy. No severe adverse events were reported. Twelve month post-treatment significant improvement in motor function was seen in all patients and follow up for up to four years after treatment showed long-term tolerability. However these results need to be treated with caution as no placebo group was included in this trial and all positive effects of this study were within the placebo range that has been reported in other clinical trials for Parkinson's disease using surgical techniques (Palfi et al., 2014).

1.4.3. Development of Lentiviral Vectors

In the early 1990s studies on HIV-1 viruses were carried out in which HIV-1 produced from viral genes expressed from two separate DNA plasmids, one plasmid encoding HIV-1 provirus with a deletion in *env* and a second DNA plasmid encoding Env protein. These early vectors could transfer non-viral genes that were inserted into the proviral plasmid. The viral *nef* gene was replaced by the chloramphenicol acetyltransferase (CAT) gene conferring drug resistance to the target cell (Helseth et al., 1990). Since then a range of other improvements have been designed making LVs not only more efficient and safer but allowing them to be used in a wider range of therapeutic applications, as described below.

The first generation of lentiviral vectors split the viral genome into three plasmids. The packaging plasmid contains *gag-pol* and regulatory as well as accessory genes expressed from a strong viral promoter. V*pu* and *env* genes were deleted. HIV-1 gp120 was replaced by VSV-G or amphotropic MLV Env and expressed from a separate plasmid. The 3'LTR in the packaging and Env plasmids were replaced by the polyadenylation site of the insulin gene. The transfer vector contained, as opposed to the other two plasmids, a packaging signal as well as both LTRs and RRE, the latter for increased export of mRNA mediated by Rev binding to RRE. Only the transgene is expressed from this construct and no viral genes (Naldini et al., 1996). In the second generation LV system most accessory genes, including *vif*, *vpr*, *vpu* and *nef* were

deleted from the packaging construct as they were shown not to be necessary for vector production in 293T cells (Zufferey et al., 1997, Kim et al., 1998). In the third generation LV packaging system *tat* was removed and *rev* provided from a separate, fourth, DNA plasmid. Tat independence was achieved by replacing the enhancer and promoter sequences in the 5'U3 region with a strong viral promoter, such as the RSV or CMV promoter. This increased vector safety considerably by reducing the number of viral genes to three and lowering the likelihood of generation of replication-competent HIV-1 like viruses as for this at least three recombination events were necessary (Dull et al., 1998, Kim et al., 1998).

Vector production and performance can be improved by further modifications of the DNA sequences used in vector packaging systems. Transduction efficiency can be increased when the central polypurine tract (cPPT) derived from pol is included into the transfer vector upstream of the transgene expression cassette and it was suggested to have a positive influence on vector nuclear translocation (Follenzi et al., 2000). Insertion of the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) downstream of the transgene in HIV-1 as well as MLV-derived vectors increases transgene expression by increasing the amount of unspliced RNA in target cells. The exact mode of action of WPRE has not been fully elucidated but is likely to be at the posttranscriptional stage such as nuclear export of viral mRNAs (Zufferey et al., 1999). The HIV-1 genome contains AU rich sequences that destabilise the viral mRNA and was first found in HIV-1 *gag* (Schwartz et al., 1992). Codon optimisation of HIV-1 Gag-Pol eliminates AU-rich sequences increasing Gag-Pol expression, resulting in a Rev independent packaging system (Kotsopoulou et al., 2000).

A problem for the efficiency of gene therapy is the silencing of provirus expression caused by epigenetic changes, such as CpG methylation, rather than mutations in the vector genome DNA (Ellis, 2005). It was suggested that vector silencing can either occur immediately after integration and takes place in the majority of proviruses or occurs long term after stable vector expression (Mok et al., 2007). To circumvent potential silencing events insulators have been used. An insulator is a DNA sequence that is able to block the activation of a promoter by an

enhancer when placed between the two, preventing gene activation or silencing as described for the 250 bp sequence of the 5' end of the chicken hypersensitive site-4 (cHS4) insulator. To avoid effects of the enhancers around the LV integration site on transgene expression an insulator has been introduced into the 3'LTR of the transfer vector increasing transgene expression (Chung et al., 1997). A minimal sequence of 400 bp of the 3'cHS4 sequence was shown to increase transgene expression levels by two-fold in murine hematopoietic stem cells (HSCs) compared to vectors without insulators (Arumugam et al., 2007). Aside from the problem that HS4 integration into transfer vectors reduces vector titers, in human HSCs only minimal effects on transgene expression could be seen (Uchida et al., 2011). Nevertheless, minimal HS4 sequences have been used in LVs for transduction of CD34+ cells in the ß-thalassemia-trial. No dominant clones were present after long term follow up (Leboulch, 2013) although there is no evidence that this absence is due to the barrier function of HS4 insulators. In another study a 650 bp HS4 sequence (a combination of the 250 and 400 bp sequence described above) and synthetic scaffold attachment regions (SARs) combined in the 3'LTR of SIN-LVs have been tested. These insulated LVs were used to transduce human embryonic stem cells (hESCs) and increased and prolonged transgene expression as well as reduced expression variability compared to LVs without the chimeric insulator were seen with minimal effects on vector titers (Benabdellah et al., 2014).

Locus control regions (LCRs) and ubiquitously acting chromatin-opening elements (UCOEs) are genetic regulatory elements that have been exploited to increase transgene expression in LVs. The LCR within the human ß-globin gene (*HBB*) specifically regulates the *HBB* family. In LVs for the treatment of ß-thalassemia, ßLCR segments have been used to enhance gene expression of the mini-*HBB* gene (May et al., 2000). UCOEs were identified in the human housekeeping HNRPA2B1-CBX3 locus and consist of 5' dual divergently transcribed promoters separated by a methylation-free CpG island. A2UCOEs, derived from the human HNRPA2B1-CBX3 locus, can be employed to drive expression of transgenes and have been shown to be superior to SFFV- or CMV promoter driven GFP expression within cells of peripheral blood cell lineages as well as

bone marrow derived HSCs after *ex vivo* transduction and transplantation in mice. Furthermore it was shown that A2UCOEs can express therapeutic levels of the *IL2RG* gene delivered by *ex vivo* transduction of HSCs and transplantation in the mouse X-linked SCID model (Zhang et al., 2007). A follow-up study documented that A2UCOEs can prevent transgene silencing in murine pluripotent embryonic carcinoma stem cells up to 44 days compared to SFFV-driven expression which was silenced after 17 days (Zhang et al., 2010). In a recent study aberrant splicing of LV-initiated transcripts driven from A2UCOE has been described that can potentially cause insertional mutagenesis. Clones were selected in which LVs, driving GFP from A2UCOE, integrated into the cellular growth hormone receptor gene (*Ghr*) in order to study effects of A2UCOE on the expression of the neighbouring gene *Ghr*. Fusion mRNAs of A2UCOE and *Ghr* were detected and were shown to be derived from the native or cryptic splice donor sites within the A2UCOE and *Ghr*. Point mutations in the A2UCOE splice donor sites could abolish the aberrant splicing potentially preventing this insertional mutagenesis effect (Knight et al., 2012).

To reduce the risk of insertional mutagenesis integration-deficient lentiviral vectors (IDLVs) have been developed. Typically they have mutations in the integrase gene (*IN*) with mutations in the catalytic core residues of *IN* gene resulting in integration deficient viral DNA (Leavitt et al., 1996). Mutations in the 5'LTR-U3 and 3'LTR-U5 attachment sites have also resulted in integration deficient vectors (Masuda et al., 1998). A study using IDLVs showed that these vectors can transduce post-mitotic cells such as retinal pigment epithelium (RPE) and neuronal cells in rodents *in vivo* and showed similar transduction efficiencies compared to integrase-proficient vectors. Furthermore functional rescue in an *in vivo* mouse model of retinitis pigmentosa was achieved by subretinal injection of IDLV making them attractive for application in clinical trials (Yanez-Munoz et al., 2006). Zinc-finger nucleases (ZFNs) are hybrid restriction enzymes linking a cleavage domain (*FokI*) to a designed zinc-finger protein (ZFP). ZFN heterodimers, made from two different ZFNs, can be used for site specific double-strand break by binding to the target site followed by homology-directed repair (HDR) through the

presence of a donor sequence delivered to the cell alongside two different ZFNs (Miller et al., 2007). IDLVs have been used to deliver ZFNs and a donor DNA sequence in the same vector. The donor sequence contains a silent point mutation resulting in the insertion of a novel restriction site and correction of the endogenous sequence with the donor-encoded sequence by HDR. This was shown in different human including haematopoietic progenitors targeting the human IL-2 receptor common gamma-chain (*IL2RG*) locus. This study proved that delivery and correction of a transgene in a specific genomic site of non-dividing cells can be achieved eliminating the risk of insertional mutagenesis (Lombardo et al., 2007). Potential off-target ZFN activity and long term re-population and differentiation studies will be required to assess its safety.

So far, no replication competent lentiviruses (RCLs) have been reported from using second or third generation packaging systems, however there is a risk of formation of replication RCLs caused by homologous recombination of sequences that are shared in transfer and packaging plasmid (Sakuma et al., 2012). One of these sequences is the encapsidation signal extending from Ψ into the 5' part of *gag.* The first 40 nucleotides of *gag* are introduced upstream of the transgene expression cassette in transfer vectors for its optimal packaging (Luban and Goff, 1994). Other homologous sequences in transfer and packaging plasmids include cPPT derived from *pol* and RRE, both being potential sites for homologous recombination.

Other developments in lentiviral vector design widen the range of their potential clinical application. The concept of using LVs as a vaccine has been reported in mice challenged with an Ova-tumour cell line. LVs transduced murine dendritic cells (DCs) deliver the SFFV driven OVA antigen (Ova-LVs). Targeting the MHC II presentation pathway the transduced DCs were shown to stimulate Ova-specific $CD8^+$ T cells. Ova-LVs could also activate Ova-specific $CD4^+$ and $CD8^+$ T cells after direct injection (Rowe et al., 2006). Furthermore maturation of DCs through NF-κB activation in mice could be induced by LVs delivering the adjuvant vFLIP from kaposi's sarcoma-associated herpesvirus (Rowe et al., 2009). Using lentiviral vectors as a vaccine means vectors will be applied in larger group of patients than in currently treated small cohorts of

patients with comparably rare diseases. To meet supply demands large scale vector production needs to be established. In search of a non-toxic alternative to VSV-G Env (see section 1.4.5) other envelope proteins need to be found. Lentiviral vectors pseudotyped with the envelope protein of alphavirus Ross river virus (RRV-LVs) were characterised in terms of their suitability to infect murine and human DCs *in vitro* and immunise mice *in vivo*. RRV-LVs could transduce murine and human DCs as well as stimulate human T cells, however a 50 fold higher dose of RRV-LVs than of LVs pseudotyped with VSV-G was required to generate a comparable T cell response (Lopes et al., 2011). Other envelope proteins for pseudotyping of LVs are described in section 1.4.5.

1.4.4. Non-HIV-1 Lentiviral Vectors

Efficient gene transfer into dividing and non-dividing cells of HIV-2- (Poeschla et al., 1998) and simian immunodeficiency virus (SIV)-derived vectors (Negre et al., 2000) has been reported as safer alternatives to HIV-1 vectors. HIV-2 is less pathogenic in humans. Replication and encapsidation deficient HIV-2 based vectors with a deletion in the packaging and *env* sequence as well as replacement of the 3'LTR with the bovine growth hormone poly A sequence were produced. It was shown that these vectors could transduce non-dividing cells such as terminally differentiated human macrophages and human neuronal cell culture model neurons (NTN2) but did not show ß-galactosidase or GFP transgene expression in human CD34+ CD38- hematopoietic progenitor cells (Poeschla et al., 1998). VSV-G pseudotyped-SIV-based LVs were constructed that were comparable to a third-generation HIV-1 packaging system, being *tat*-independent, devoid of *env* accessory genes *vif*, *vpx* and *vpr* as well as the 3'LTR, which was replaced with SV40 polyA. These vectors were able to transduce human 293 as well as mature and immature DCs (Negre et al., 2000). In another study gene transfer and expression of SIV-based vectors were tested on X-CGD patients' cells. SIV-derived RD114/TR pseudotyped vectors, equivalent to the vector design of third generation HIV-1 based vectors were shown to

transduce *ex vivo* CD34+ PBMCs from X-CGD patients delivering the human $gp91^{phox}$ (Naumann et al., 2007). These vectors could be a potential safer alternative to HIV-1 based LVs once their application in other diseases has been further investigated.

Further developed towards clinical applications are non-primate lentiviral vectors derived from equine infectious anaemia virus (EIAV) or feline immunodeficiency virus (FIV). Non-primate lentiviral vectors have been proposed to be a safer alternative to HIV-1 based LVs as they cannot replicate in human cells in case of the development of replication competent viruses. For example the functional receptor of FIV is feline CD134 on $CD4^+$ T cells but not human CD134 (Shimojima et al., 2004). Non-primate lentiviral vectors have been developed including EIAV-based vectors pseudotyped with various envelopes such as VSV-G or RVG (Mitrophanous et al., 1999). Preclinical safety studies of VSV-G pseudotyped EIAV-based Tat independent LVs for the treatment of age-related macular degeneration and juvenile macular dystrophy showed that these vectors were well tolerated and localised to the site of administration (Binley et al., 2012, Binley et al., 2013). Recently published results of the first clinical phase 1/2 trial have been published using VSV-G pseudotyped EIAV-based vectors for the treatment of Parkinson's disease (Palfi et al., 2014) (see section 1.4.2.7).

1.4.5. Viral Vector Pseudotyping

Pseudotyped lentiviral vectors, like other retroviral vectors, have been developed in combination with heterologous envelope proteins. Envelope proteins from different viruses recognises each a specific cellular receptor hence the pseudotype of a retroviral vector determines the cell type it can transduce, also called the vector tropism. The vesicular stomatitis virus glycoprotein (VSV-G) has been the most commonly used envelope for lentiviral vectors due to its high stability, wide host range and cell tropism (see section 1.2.14). Disadvantages of VSV-G for the use in vector pseudotyping were reported in comparison with other envelopes including amphotropic murine leukaemia virus (MLV-A) Env, a modified

gibbon ape leukaemia virus (GALV+) Env and two modified RD114 envelope proteins. VSV-G Env was sensitive to inactivation by human complement and of lower resistance to freeze-thawing when harvested in serum-free media (Strang et al., 2004). VSV-G is also known to be cytotoxic at high concentrations and during long-term expression, hence not suitable for the use in a stable producer cell line (Sakuma et al., 2012).

Targeting a specific cell type can be achieved by using alternative envelope proteins. Relander et al. (2005) tested pseudotyped lentiviral vectors for the transduction of haematopoietic stem cells (HSC) with various gammaretroviral envelopes including amphotropic MLV 4070 A, the modified envelope RDpro derived from feline endogenous virus RD114, and the GALV Env with an MLV cytoplasmic tail (GALV+) produced by the STAR cell line. Vectors bearing RDpro gave the best transduction efficiency in CD34+ cells compared to the other retroviral envelopes. RDpro is an improved form of the RD114 envelope, when used in vector packaging, producing higher titers of vectors compared to its wild type form. To create RDpro the R peptide cleavage site was replaced by a matrix-capsid cleavage site from HIV-1 Gag (Ikeda et al. 2003) which was predicted to increase R peptide cleavage by HIV-1 protease, thought to be a crucial condition for viral host cell entry.

LVs incorporating Edmonston measles virus (MV) glycoproteins haemagglutinin (H) and fusion protein (F) (H/F-LVs) could transduce completely resting B (Frecha et al., 2009) and T cells (Frecha et al., 2008) without prior stimulation. Certain envelope proteins can also alter intracellular behaviour of viral vectors. Pseudotyping EIAV derived vectors with rabies virus glycoprotein (RVG) allows retrograde transport of vectors along motor neuron cell axons (Mazarakis et al., 2001). Transduction efficiency could be improved by about 25-fold in using a chimeric RVG-VSV-G envelope in comparison to RVG pseudotyped HIV-1 derived LV. The chimeric envelope is comprised of the RVG receptor-binding domain by replacing the cytoplasmic, transmembrane and extracellular stalk regions of RVG with the corresponding regions from VSV-G (Carpentier et al., 2012).

1.5. Viral Vector Production: Transient Transfection

For the production of lentiviral vectors 293T cells are most commonly used as they are highly transfectable (Pear et al., 1993). They are derived from human embryonic kidney (HEK) cells transformed with sheared adenovirus type 5 DNA (Graham et al., 1977). The modified form 293T cells, stably expressing the SV40 T-large antigen (T-Ag) facilitate the replication of plasmids containing a SV40 replication origin. Depending on which viral vector generation is used transient transfection involves adding DNA plasmids at a defined ratio for most efficient transfection. Plasmids include the transfer vector, packaging and envelope constructs as well as plasmids for regulatory genes (in third generation systems) to the packaging cells in combination with a transfection reagent such as calcium-phosphate or lipid-based reagents.

Traditionally calcium-phosphate precipitation (Ca_2PO_4) has been used during which precipitates of calcium-phosphate bind to DNA plasmids. However, this method has several disadvantages as the transfection efficiency can vary depending on factors such as the density and general condition of the cells influencing transfection efficiency. Factors easier to control during transfection with Ca_2PO_4 precipitation is the formation of calcium phosphate precipitates. The concentration of calcium and phosphate as well as the pH of the transfection buffer affects the size of formed precipitates which in turn defines their capacity to bind DNA. Precipitate stability after addition to the culture medium is dependent on the pH of the culture medium (Jordan and Wurm, 2004).

Alternatively to Ca_2PO_4 precipitation the polyethylenimine (PEI) based transfection method can be used and gives higher yields of infectious particles compared to Ca_2PO_4 precipitation (Segura et al., 2007, Toledo et al., 2009) but optimisation of PEI ratio to DNA amounts is required in order to form PEI-DNA complexes (Reed et al., 2006). Commercially available lipid-based transfection reagents, such as Fugene®, have been used in lentiviral vector production (Kosaka et al., 2004) and allow reproducible and efficient transfections as they are less pH dependent. They also require less than one-third of the DNA amounts compared to the Ca_2PO_4

method (Coleman et al., 2003). Transient transfection results in temporary or transient expression of viral genes and vector particles are normally produced over a short period of five to six days before cells begin to detach and die (Segura et al., 2013). Cell factory systems (Nunc, ThermoScientific) or HYPERFlask^{TM} (Corning) have been used to scale up vector production. Alternatively transient transfection using PEI in a suspension-growing clone of 293 cells (293E) can produce lentiviral vectors in bioreactors with a capacity of 3 litre (Segura et al., 2007). This method however, requires large amounts of reagents as well as DNA plasmids and consistency between production batches is less easily to control than in stable producer cells. On the other hand transient transfection provides flexibility in vector particle construction avoiding the development of a stable producer cell line which can take several months.

1.6. Stable Vector Production Systems

Currently most lentiviral vectors used in clinical trials are produced by transient transfection. However, cell lines that have viral components stably integrated into their genome, stable producer cells, would be preferable for large scale vector production. Large batches of vectors with high reproducibility and of consistent quality at low costs can be generated. The use of virus producing cell lines (VPCLs) or packaging cell lines (PCLs) can avoid: 1) the risk of recombination between transfected plasmids, 2) medium contamination with DNA plasmids that can co-purify with vectors and need to be removed during downstream processing and 3) variability between batches in transient transfection. Producer cell lines stably express *trans*-acting vector elements including Gag-Pol, Env and, if the third generation LV production system is used, Rev.

Production of high amounts of infectious vector particles for clinical trials by a PCL has shown to be difficult as a sufficient level of constitutive HIV *gag-pol* expression is hampered due to its toxicity to producer cells (Kräusslich 1992). Other components such as Rev (Miyazaki et al. 1995) and vesicular stomatitis virus-G protein (VSV-G) (Burns et al. 1993) have also been

shown to be cytotoxic to producer cells if over-expressed. This promoted the development of cell lines with inducible expression of viral elements, for example tetracycline-regulated promoters. In the Tet-off system the expression of more than one viral gene can be controlled. In a study using this system the vector genome was transduced with transiently produced LVs (Klages et al., 2000). 293G cells express tetracycline transactivator protein (tTA) constitutively as well as VSV-G Env under the control of a tetracycline-inducible promoter. To allow indirect controlled expression of Gag-Pol, 293G cells were stably co-transfected with HIV-1 Gag-Pol and Rev. Rev expression is driven by the tet operator (tet O) fused to a minimal CMV promoter forming the tetracycline response element (TRE). In the absence of tetracycline or its derivative doxycycline tTA binds to tetO activating transcription of Rev. Rev expression in turn controls the expression of Gag-Pol since Gag-Pol transcripts are exported from the nucleus only in the presence of Rev and otherwise degraded in the nucleus. Hence, removal of tetracycline (or doxycycline) from the cell culture medium induced Gag-Pol expression. The Tet-off system for expression of Gag-Pol as well as VSV-G Env by tetracycline induction was also used in another PCL with introduction of the transfer vector by several rounds of infection (Farson et al., 2001).

A disadvantage of inducible systems is that they require the removal of tetracycline or doxycycline from the culture medium which is not practical in large scale vector production. Other inducible transcription systems, like the Tet-on and the cumate switch, have been applied that require addition of an inducer to the growth medium (Broussau et al. 2008). In this stable producer cell line regulation of VSV-G *env* is achieved by using a Tet-on system that requires the addition of doxycycline (Dox) for expression. *Rev* expression is regulated by using the cumate switch inducible system derived from the *p-cym* operon of *Pseudomonas putida* based on the addition of the inducer cumate to the culture medium. A disadvantage of this system is that Dox and cumate remain in the vector containing cell supernatant and their removal during downstream vector purification is required. On the other hand this PCL can also grow in medium without the addition of serum and was cultivated in suspension for large

scale vector production. Avoiding serum in the vector production systems has the advantage of cost reduction and removal of fewer impurities during vector purification. An inducible lentiviral packaging cell line based on equine infectious anemia virus (EIAV) for Parkinson's disease was developed successfully based on VSV-G *env* and EIAV *gag-pol* expression induction by the addition of Dox to the cell culture medium producing vectors with stable titer for 7 weeks (Stewart et al., 2009) and up to 16 weeks (Stewart et al., 2011). None of the inducible lentiviral vector producer cell lines have been used in clinical trials so far. Transient transfection allowed production of adequate amounts of vectors for the treatment of small patient groups. In other cell lines shortcomings of the induction system have been documented including high basal expression of uninduced promoter and insufficient levels of activation of the inducible promoter (Ni et al., 2005).

Using an alternative approach for stably introducing vector packaging components the continuous HIV-1 packaging cell line STAR has been developed in the laboratory at UCL (Ikeda et al. 2003). In 293T cells high expression of HIV-1 Gag-Pol was achieved by introducing a codon optimised *gag-pol* expression cassette (Kotsopoulou et al. 2000) into the cells by retroviral vector transduction and selection of a high expression locus. Stable expression *tat* and *rev*, was achieved by the same method, transduction with MLV gammaretroviral vectors with full 3'LTR and internal CMV promoters driving expression of the viral gene. As described in section 1.4.5, VSV-G Env is not suitable for stable expression in a producer cell line. For selection of an alternative viral envelope several gammaretroviral envelopes were compared, including amphotrophic MLV 4070A (Ampho), RDpro and GALV+. GALV+ is a GALV Env with an MLV cytoplasmic tail and RDpro a modified form of RD114 glycoprotein with the R cleavage site replaced by the HIV-1 protease cleavage site (Ikeda et al., 2003) increasing efficiency of envelope processing during maturation of vector particles. Titers of up to $8.5x10^6$ transducing units per ml could be achieved when transducing STAR-RDpro cells with the non-self-inactivating (non-SIN) vector genome HV delivered by a HIV-1-VSV-G pseudotyped vector, creating STAR-RDpro-HV cells. Among vectors pseudotyped with different envelope proteins

GALV+ or Ampho MLV vectors pseudotyped with RDpro Env showed the best transduction efficiency. In comparison to transiently produced vector particles pseudotyped with VSV-G Env, STAR-RDpro produced less infectious particles per p24 ng, however titers are still sufficiently high for producing lentiviral vectors for clinical trials. STAR-RDpro cells require further improvement before they can be used for clinical grade production. The two main problems with the current design are 1) the 293T clone used for production of these cells is non-traceable and 2) HIV *gag-pol* is packaged into virions (Ikeda et al., 2003). Currently an improved form of STAR cells is under development considering these aspects. Generally, the development of STAR-RDpro has shown that LVs can be stably produced without the need for an inducible system for HIV-1 *gag-pol* or viral envelope expression.

The above described cell lines generate non-SIN LVs are not suitable for many clinical applications (see section 1.4.3). Transduction of a SIN-vector genome into the producer cell genome is not applicable since the duplication of the deletion in the 3'LTR region during reverse transcription would leave the integrated provirus promoterless preventing the expression and packaging of the vector genome into vector particles. Transient transfection of a SIN-vector genome does result in lower vector titers compared to vectors produced by transduction of a non-SIN-vector genome (Ikeda et al., 2003). To circumvent this problem a conditional SIN (cSIN) vector was designed. In cSIN vectors the 3'U3 region enhancer and promoter sequence is deleted and replaced by a tetracycline response element (TRE) initiating transcription of the vector genome RNA upon tTa binding. However improvement of the cSIN vector genome sequence is required as a basal expression of the full vector genome RNA was detected even in the absence of tTa in target cells (Xu et al., 2001).

To achieve high transgene expression the vector genome can also be introduced by site specific integration. This system allows the stable integration of the SIN vector genome construct into a predefined locus. Using the FRT-flp system, a high expression locus in the producer cells is tagged followed by RMCE with a targeting vector (Schucht et al., 2006). This is

achieved by co-transfection of the Flp-tagged cells with a targeting vector encompassed by FRT sites (the SIN-vector genome) and a DNA plasmid expressing the tagging site specific recombinase flipase.

In the LV packaging cell line GPRG-TL20-GFP, the introduction of the SIN vector genome (see section 1.4.3) into producer cells was achieved by concatemeric array transfection technique (Throm et al., 2009). Vector components, Gag-Pol, Rev and Env were introduced sequentially by transduction with γ-RVs, in which *rev* expression was controlled by the tet-off system. Finally, the vector genome was transfected using a concatemeric array. These arrays are two linked vector genome expression cassettes and are regulated by doxycycline allowing increased transgene expression in combination with the safety of a SIN vector genome design. The concatemeric array transfection technique was also used by others confirming that this system can produce SIN-LVs of high titers with ~10^7 transducing units per ml cell supernatant (Lee et al., 2012).

1.7. Aim of Thesis

The project outlined in this thesis was aimed to identify possible protein-protein interactions between host and lentiviral vectors (LVs) in an attempt to elucidate which of these proteins might be important for vector particle production, focusing on particle assembly and budding from the producer cell. LVs are being used for the treatment of relatively rare diseases, including primary immunodeficiencies and neurodegenerative diseases but also increasingly for applications such as cancer therapy that will target diseases with a higher prevalence. Furthermore promising results of early clinical trial stages will allow the use of LVs in larger cohorts of patients in clinical trial stages III and IV. This development is leading to an increased demand in vector material and requires improvement in vector production methods to produce large batches of vectors with high infectious titers. This in turn would benefit from a good understanding of vector particle formation within the producer cell.

Characterisation of cellular protein interactions with viral proteins, in particular identifying interactions that promote production of functional particles, could potentially help advance viral vector production. This investigation set out to contribute to a better understanding of vector biology. LV samples were produced by the transient transfection method, currently used for vector preparation in clinical trials, pseudotyped with the widely applied VSV-G. Alternatively stably produced RDpro-pseudotyped vectors were prepared by the packaging cell line STAR-RDpro. Following vector purification by size exclusion chromatography (SEC), subsequent protein content analysis by liquid-chromatography mass spectrometry (LC-MS/MS) was applied in order to find differences and similarities in host cell protein composition between vector samples. To identify proteins that might be involved in vector assembly in 293T or 293T-derived STAR-RDpro cells selected proteins were knocked down in producer cells by siRNA for further analysis of their function on vector production.

2. Materials and Methods

2.1. Cell Culture

HEK293T (293T) cells are derived from 293 human embryonic kidney cells and express the SV40 large T-antigen (Graham et al., 1977). The stable producer cell line, STAR-RDpro-pHV (Ikeda et al., 2003), is derived from 293T cells by transduction with MLV vectors delivering the codon optimised HIV-1 *gag/pol* (Kotsopoulou et al., 2000), *tat* and *rev* each driven by a cytomegalovirus (CMV) promoter and stably transfected with the gammaretroviral envelope RDpro. The vector genome encoding a green fluorescent protein (GFP) expression cassette driven by the spleen focus-forming virus (SFFV) promoter and is stably integrated in the STAR-RDpro-pHV cells after transduction with a HIV-1 derived vector (Ikeda et al., 2003). For stable production of LVs without genomic RNA, the STAR cell line STAR-RDpro is used which does not contain a vector genome. All cell lines were maintained in Dulbecco's modified Eagle medium (DMEM) (Sigma-Aldrich, Gillingham, UK) completed by the addition of 10% fetal calf serum (Appleton Woods, Birmingham, UK), 2 mM L-glutamine, 100 units/ml penicillin, 100 µg/ml streptomycin (all by Sigma-Aldrich). Cells were usually passaged every 48 or 72 hours at a ratio of 1:10 or 1:20, respectively.

2.2. Amplification and Purification of DNA plasmids

2.2.1. Preparation of Chemically Competent Bacteria Cells

Chemically competent *Escherichia coli* (*E.coli)* XL1-Blue (Stratagene, Santa Clara, CA, USA) were prepared for transformation with DNA plasmids. Bacteria were streaked on an LB agar plate containing 12.5 µg/ml tetracycline (both by Sigma-Aldrich) and incubated overnight at 37°C. The next day a single colony was picked and grown overnight in 4 ml Luria Bertani (LB) medium (Lennox) (Sigma-Aldrich) with 12.5 µg/ml tetracycline. After 8 hours incubation on a shaker

(300 rpm) at 37°C 100 ml of LB medium (without antibiotics) were inoculated with 1 ml of the overnight culture and incubated on a shaker at 37°C for approximately 2.5 hours or until OD_{600} was between 0.3-0.6. After cooling of the culture on ice for 5 minutes and centrifugation at 6000 x g for 10 minutes (4°C) the bacteria pellet was resuspended in 50 ml TBF-1 (30 mM of potassium acetate, 100 mM potassium chloride, 10 mM calcium chloride, 50 mM manganese chloride, 15% glycerol; all by Sigma-Aldrich). The suspension was incubated on ice for 5 minutes and centrifuged at 500 x g for 10 minutes (4°C) followed by re-suspension in 4 ml TBF-2 (10 mM MOPS, 75 mM calcium chloride, 10 mM potassium chloride, 15% glycerol; all supplied by Sigma-Aldrich) and incubation for 15 minutes on ice. Aliquots of 50 µl were prepared on ice and stored at 80°C.

2.2.2. Transformation of Competent Cells with Plasmid DNA

Transformation efficiency of the competent XL1-Blue *E.coli* was determined. An aliquot of cells was thawed on ice and 0.1 ng of control DNA (1 µl of 0.1 ng/µl serially diluted pHV plasmid DNA) added. After incubation for 20 minutes on ice cells were transferred to a water bath at 42°C and incubated for 2 min. This was followed by incubation for 2 minutes on ice and then bacteria were mixed with 950 µl of LB medium (without antibiotics) and incubated at 37°C on a shaker for 1 hour. 10 µl of the culture were diluted in 990 µl of LB medium and 100 µl spread on an LB agar plate with 100 µg/ml of ampicillin (Sigma-Aldrich) and incubated overnight at 37°C. To calculate the transformation efficiency colonies on the plate were counted. For example if 100 colonies were counted, the transformation efficiency is: number of colonies on plate/ ng of DNA plated x 1000 ng/µg, hence 100 transformants/0.0001 µg x 1000 ng/µg = 1 x 10^9 transformants/µg.

For transformation of competent cells with DNA 10 to 50 ng of plasmid DNA were added to the thawed cells and the above described procedure was followed, with the difference that 200 µl

of the 1 ml culture were spread on the agar plate before incubation of the plate at 37°C overnight.

2.2.3. DNA Plasmid Preparation

Plasmid DNA was purified from transformed *E.coli* using a DNA plasmid kit (Qiagen, Limburg, Netherlands) following the manufacturer's protocol. A single colony of transformed bacteria from a LB agar plate was picked and used to inoculate 5 ml of LB medium containing the respective antibiotic. After 8 hours shaking incubation at 37°C the required volume of LB medium containing the respective antibiotics were inoculated at a dilution of 1:500 with the starter culture and incubated on a shaker overnight at 37°C. Following plasmid DNA extraction by alkaline lysis of bacteria and plasmid DNA purification, DNA pellets were resuspended in 1 to 2 ml of TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA, supplied in kit). Plasmid DNA concentration was quantified using the NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, Hemel Hampstead, Herts, UK) measuring the absorbance A_{260} and aliquots kept at -20°C or -80°C for long term storage.

2.2.4. DNA Plasmids

DNA plasmids were used to make transiently produced LVs (chapter 3) and GIPZ-LVs (chapter 4). Transient transfection is described in sections 2.3.1 and 2.5.1. Plasmid pHV is the third generation lentiviral genome vector. The 3' and 5' long terminal repeats (LTRs) of pHV were derived from pH7G (provided by Oxford Biomedica, Oxford, UK; (Ikeda et al., 2003). The U3 region, containing enhancer and transcription regulatory signals, of the 5'LTR was replaced by a CMV promoter. Internally, pHV encodes enhanced GFP (eGFP) under the control of the SFFV promoter and the packaging sequence psi (Ψ), central polypurine tract (cPPT) upstream as well as the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) downstream of eGFP (Ikeda et al., 2003). The transfer vector pRRLSIN.cPPT.PGK-GFP.WPRE is a self-

inactivating (SIN) vector genome plasmid with a deletion in the U3 of the 3'LTR encoding eGFP driven by the promoter of the human phosphoglycerate kinase gene (PGK), and contains cPPT, Ψ as well as WPRE. Plasmids p8.91 and pMDLg/pRRE contain wild type HIV-1 *gag-pol* under the control of the CMV promoter. Plasmids pMDG and pMDG.2 contain VSV-G envelope under the control of a CMV promoter. Plasmid pRSV-Rev encodes Rev driven by rous sarcoma virus (RSV) U3. The DNA plasmid RDpro encodes the RD114 envelope, where the R-peptide cleavage site is replaced with a HIV-1 protease cleavage site, driven by MLV LTR (Ikeda et al., 2003). The DNA plasmid pGIPZ is a SIN-LV genome containing the elements of cPPT, Ψ and WPRE with a shRNA as transgene. The shRNA is based on the endogenous miRNA 30 (mir-30). To mark shRNA expression the reporter gene turbo GFP (tGFP) driven by the CMV promoter is included and placed upstream of an internal ribosome entry site (IRES) that drives the expression of an antibiotic selection marker, puromycin as well as expression of the shRNA (Thermo Fisher Scientific). An overview of all plasmids including their source is given in the appendix (Table 2 and Figure 4).

2.3. Lentiviral Vector Preparation and Characterisation

2.3.1. Vector Production

Transiently produced LVs (in chapter 3 and Table 1) were prepared by calcium-phosphate four-plasmid co-transfection of 293T cells using third generation HIV-1 vector plasmids (adopted method by (Kutner et al., 2009). On the day before the transfection, $8x10^6$ 293T cells were plated per 15 cm dish in 18 ml complete DMEM. The DNA plasmids used to transfect cells of one 15 cm plate were 7 µg pMD2.G encoding VSV-G, 7 µg pRSV-Rev encoding Rev, 14 µg pMDLg/pRRE encoding HIV-1 Gag-Pol and 21 µg pRRLSIN.cPPT.PGK-GFP.WPRE encoding a SIN-LV genome including the marker protein eGFP driven by a PGK promoter; representing a ratio of 1:1:2:3. The day after transfection, the supernatant was changed to 18 ml fresh complete DMEM per plate, followed by the first harvest 48 hours post-transfection. For the second

harvest, fresh DMEM was added to the cells and collected 72 hours post-transfection (day 4 and 5 after plating). First and second supernatants were filtered through 0.45 µm cellulose acetate (Schleicher & Schwell Micro Science, NJ, USA) pooled and centrifuged at 19000 rpm for 2 hours at 4°C (Beckman Coulter, L-90K, SW28). Vector supernatants from one 15 cm dish were concentrated from 18 ml to 900 µl. For preparations of vector samples up to batch #54 pellets were soaked for 1 hour in complete DMEM at 4°C and resuspended to a final concentration factor of 40. Pellets of vector samples from batch #55 onwards, including those that were analysed by LC-MS/MS analysis, were soaked for 1 hour in Opti-MEM® Reduced Serum Medium, GlutaMAX™ (Invitrogen, Carlsbad, CA, USA) at 4°C and resuspended to a final concentration factor of 40. Aliquots were stored at -80°C. The sample at this stage is referred to as 'crude sample'.

For lentiviral vector production by stable producer cells, STAR-RDpro-pHV cells (Ikeda et al., 2003) were seeded at a density of $8x10^6$ cells per 15 cm dish. After 48 hours the medium was changed to 18 ml complete DMEM per plate for the first harvest. This was followed by the first and second harvest 24 hours and 48 hours later (day 4 and 5 after plating). Filtration of harvest was carried out as described above and LVs concentrated at a reduced speed of 8500 rpm and spin time of 1.5 hours at 4°C (Beckman Coulter, L-90K, SW28), according to (Strang et al., 2004).

2.3.2. Lentiviral Vector Purification using Size Exclusion Chromatography

For vector purification, a Gilson liquid chromatography system was used, consisting of a pump (model 306), an autosampler (model 231XL) as well as a fraction collector (model FC203B) connected to a temperature controlled rack. Sample injection was controlled by the software Trilution LC 2.1 and sample elution and detection were controlled by Unipoint 5.11 (all parts and software by Gilson, Middleton, WI, USA). Samples were separated by a cooling jacked XK16/70 column packed with Sephacryl-500–HR medium (both by GE Healthcare, Little

Chalfont, UK). Sephacryl-500–HR medium consists of cross-linked copolymer beads of allyl dextran and N,N'-methylene bisacrylamide with an average particle size of 47 µm, allowing the separation of macromolecules in the range of $4x10^4$ to $2x10^7$ relative molecular mass (M_r). TEN buffer containing 150 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl pH 7.4 was used as sample running buffer (Sigma-Aldrich).

For size exclusion chromatography, 900 µl of crude vector sample were purified per column at a flow rate of 0.8 ml/minutes with a bed volume of 140 ml for a run time of 350 min, allowing the flow through of buffer equivalent to two column volumes. One 4ml fraction, the void peak fraction, was collected per run. Three columns were used in this project, depending on the exact volume of medium in each column, the elution time ranged from 64 to 68 min. To set up a newly packed column for SEC, the elution time of the void peak was determined and the collection start of the first fraction was set. Fractions were collected in a cooled rack and stored at -80°C for further processing.

2.3.3. Purified Lentiviral Vector Sample Preparation for Mass Spectrometry

To prepare SEC purified viral vectors for LC-MS/MS analysis, 4 ml void peak fractions from a total of 8 runs of 900 µl crude vector samples were pooled and dialysed to exchange the sample buffer TEN to 10 mM ammonium bicarbonate (ABC; Sigma-Aldrich), pH 8.0.

Slide-A-Lyzer dialysis cassettes (3 to 12 ml capacity, molecular cut-off 10 000 Da, Thermo Fisher Scientific) were hydrated for 2 minutes in 10 mM ABC and filled with 2x 4ml SEC void peak fractions through one of the cassettes injection ports using a 10 ml syringe and 18 Gauge needle. Samples were dialysed overnight in 3 litres of 10 mM ABC, pH 8.0. Dialysed samples were freeze-dried using an Edwards E2M2 High Vacuum Pump. Lyophilised samples were resuspended in distilled water and stored at -80°C.

2.3.3.1. Purified Lentiviral Vector Quantification using Bradford Protein Assay

For total protein quantitation samples were measured in triplicate by mixing 5 µl of sample (in 10 mM ABC; pH 8.0), 10 µl of HCl (0.1 N; Sigma-Aldrich) and 785 µl dH2O. A standard curve was generated using BSA protein standard (Thermo Fisher Scientific) in the range of 0.125-0.875 µg/ml, in duplicates. After 5 minutes incubation at room temperature with 200 µl of concentrated Bradford dye (BioRad, Hercules, CA, USA), absorbance was measured using Lambda 800 UV/Vis Spectrometer (PerkinElmer, Waltham, MA, USA) at the wavelength 595 nm.

2.3.3.2. Purified Lentiviral Vector Characterisation using LC-MS/MS

LC-MS/MS analysis was performed by Jun Wheeler, in the proteomics facility, NIBSC. In brief, purified and concentrated lentiviral vector samples were digested with trypsin in the presence of 1% Rapigest (an enzyme-compatible detergent to ensure solubilisation of membrane proteins; Waters, Milford, MA, USA) and 50 mM ammonium bicarbonate, pH 8.5 for 3 hours at 37^{o}C. HCl was added to terminate digestion and ensure breakdown of Rapigest.

LC-MS/MS was carried out using a mass spectrometry system (Thermo Fisher Scientific, Hemel Hampstead, Herts, UK) equipped with a nano-electrospray ion source and two mass detectors including linear trap (LTQ) and orbitrap, coupled with an Ultimate 3000 nano-LC system, comprising a solvent degasser, a loading pump, a nano-pump, and a thermostated autosampler. After an automated injection, the extracted peptides from each digestion were desalted in a trapping cartridge (PepMap reversed phase C18, 5 µm 100 Å, 300µ id x 5 mm length; Thermo Fisher Scientific), eluted on to a C18 reversed phase nano-column (3 µm, 100Å, 5 cm length; Thermo Fisher Scientific) followed by a 60 minutes separation under a column flow rate of 0.3 µl/minutes using linear gradient of 5 to 70% acetonitrile and 0.1% formic acid (both by Sigma-Aldrich). Separated and eluted peptides were ionised by electrospray ionisation (ESI) followed by a MS survey scan (mass-to charge-ratio, m/z 400 – 2000) in the

LTQ, sequentially selecting the five most abundant ions of peptides eluting from the LC at that time, before being passed on to the Orbi-trap. The total cycle time for each MS/MS was approximately 30 milliseconds. The Orbi-trap took accurate mass measurement with the resolution of 30,000 parts per million (ppm) and ions were then fragmented in the linear ion trap by collision induced dissociation (CID) at collision induced energy of 35%. Subsequently, fractionated ions were separated according to their m/z ratio. Data was collected in data dependent MS/MS mode with dynamic exclusion set to 2 counts. Data analysis including mass spectra processing and database searching was carried out using Thermo Proteome Discoverer 1.2. with built-in Sequest (Thermo Fisher Scientific). Initial mass tolerances for protein identification by MS were set to 10 ppm. Up to two missed tryptic cleavages were considered and methionine oxidation was set as dynamic modification. Peptide sequences by MS/MS were only included when Xcorrelation scores were greater than 1.5, 2 or 2.2 for charge states 1, 2 and 3, respectively. An unambiguous identification was considered when at least two peptides matched to the protein. The protein FASTA databases were downloaded from www.uniprot.org, release 2012-03 including the complete entries from *Homo Sapiens* (taxon identifier 9606), *Bos Taurus* (9913), human immunodeficiency virus type 1 group M subtype B (isolate HXB2) (HIV-1) (11706), vesicular stomatitis indiana virus (strain San Juan) (VSIV) (11285), RD114 virus (11834), RDpro Env (Ikeda et al., 2003) AA sequence: (Bell et al., 2010, Ikeda et al., 2003), and GFP (P42212).

2.4. Lentiviral Vector Characterisation

2.4.1. Vector Titration by Fluorescence Activated Cell Sorting (FACS)

The vector transduction titer was measured after concentration by ultracentrifugation and after SEC purification of vector particles. 293T cells were plated in 6-well dishes at a density of $2x10^5$ cells per well in 2ml complete DMEM and allowed to grow overnight. Cells were transduced with concentrated vector after serial dilution in complete DMEM containing 8

µg/ml polybrene (Sigma-Aldrich). Concentrated vector samples were serially diluted and final amounts of 2.5 µl, 0.5 µl and 0.1 µL in 1 ml complete DMEM were added to cells of each well. Five hours later the medium was taken off and 2 ml of complete DMEM added. 72 hours post transduction cells were analysed by FACS for percentage of GFP expressing cells.

To titrate SEC purified samples, 1.5 ml of a 4 ml void peak fraction was 0.2 µm filtered (cellulose acetate; Schleicher & Schwell Micro Science) and cells transduced with 400 µl, 100 µl and 50 µl of filtered SEC fractions and processed like crude samples above. After 4 to 5 hours incubation, the culture medium was exchanged with 2 ml of complete DMEM. After 72 hours the cells were analysed by FACS for percentage of GFP expressing cells.

Lentiviral vector particles carrying the pGIPZ vector genome (concentrated LV-GIPZ) used in knock-down of host cell proteins experiments were titered in a similar way. 293T cells were plated in 24 well plates at a density of $4x10^4$ cells per well in 500 µl of complete medium. The next day serial dilutions of LV-GIPZ preparations were prepared and cells transduced with final amounts of 50 µl, 10 µl and 2 µl vector sample in 250 µl complete DMEM containing 8 µg/ml polybrene per well. The medium was changed 5 hours following transduction and GFP expressing cells were counted by FACS using a BD FACSCanto™ II (BD Biosciences, San Jose, CA, USA) 72 hours after transduction.

Unconcentrated LV harvests produced by STAR-GIPZ and 293T-GIPZ cells were titered on 24 well plates with $4x10^4$ 293T cells per well. Cells were transduced with serially diluted harvests in DMEM containing 8 µg/ml polybrene; 250 µl, 50 µl and 10 µl of harvests were applied in a final amount of 250 µl per well. Five hours later the medium was changed and GFP positive cells counted by FACS 72 hours post transduction.

The number of transducing units per ml (TU/ml) was calculated using the equation: TU/ml = (no. of cells at time of transduction x [% of GFP positive cells/100])/input volume of vector sample in ml.

2.4.2. Vector Titration by p24 ELISA (Chapter 3)

SEC fractions of lentiviral vector samples were analysed by p24 ELISA (chapter 3). On day one, a 96 well plate (High Binding, Corning, by Fisher Scientific, Gillingham, UK) was prepared by coating each well with 10 µg/ml of capture antibody, affinity-purified sheep anti-HIV-p24 Gag (ARP410, CFAR, NIBSC, UK). Plates were in a closed container and left to incubate at room temperature overnight. On day two, wells were washed six times with 200 µl per well of TBS (50 mM Tris, 150 mM NaCl, pH 7.5, media facility, NIBSC), followed by blocking each well with 200 µl of 2% skimmed milk (Marvel, Original Dried Semi Skimmed Milk Powder) in TBS solution and incubated for 30 minutes at room temperature in a closed container to prevent drying out. Another wash with TBS (6x200 µl per well) followed and the p24 protein standard (ARP620, CFAR, NIBSC, UK) was added to each well in a serial dilution in phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCL and 10 mM phosphate buffer, pH 7.4, media facility, NIBSC) containing 0.1% Empigen (Millipore, Billerica, MA, USA) and 10% sheep serum (Sigma-Aldrich) (PBS/E/S), giving a p24 detection range of 78 pg to 5 ng. To 100 µl vector sample (SEC fraction) 11 µl of 1% Empigen were added, mixed, heated at 56°C for 30 minutes and added to each well. The plate was incubated overnight at room temperature in a closed container. The following day, wells were washed with TBS (6x200 µl per well) with subsequent antigen detection. The detection antibody, biotinylated monoclonal mouse anti-HIV-p24 Gag (ARP454, CFAR, NIBSC, UK), was diluted to a working concentration of 1 µg/ml in TBS containing 2% skimmed milk, 20% sheep serum and 0.5% Tween 20. 100µl diluted detection antibody were added to each well and incubated for 2 hours at room temperature. This was followed by a wash with TBS (6x200 µl per well). Extravidin (Sigma-Aldrich) was diluted 1:2000 in PBS-0.5% Tween 20, 100 µL added per well, followed by an incubation of 1 hour at room temperature. Wells were washed again and p-nitrophenyl phosphate (Sigma-Aldrich), dissolved as per manufacturer's instructions in reaction buffer (10mM ethanolamine, 0.5mM $MgCl^2$, pH 9.8; Sigma-Aldrich), added to each well in aliquots of 100 µl. The absorbance at 405 nm was recorded directly after addition of the substrate as chromogenesis starts immediately.

Readings were taken every 5 minutes for 20 minutes using an absorbance microplate reader (SpectraMax 340PC, Molecular Devices, Sunnyvale, CA, USA).

2.4.3. Vector Titration by p24 ELISA (Chapter 5)

The p24 ELISA kit 'Lenti-X p24 Rapid Titer Kit' (Clontech, Mountain View, CA, USA) was used to measure p24 protein levels in vector harvests from STAR, transiently transfected 293T cells and STAR-GIPZ and 293T-GIPZ LV producer cells presented in chapter 5. The manufacturers' protocol was followed. The assay principle is based on the binding of p24 of vectors in cell samples to anti-p24 antibodies coated on a microwell plate. This is followed by incubation with a secondary biotin-labelled antibody against HIV-1 p24. Subsequently samples are exposed to a streptavidin-horse radish peroxidase (HRP) conjugate. In the next step the soluble colorimetric substrate solution 3,3',5,5'- tetramethylbenzidine (TMB) is added and reacts with HRP to form a blue by-product whose intensity is proportional to the amount of HRP, which in turn is directly proportional to the amount of p24 in the samples. After addition of sulfuric or phosphoric acid stop solution (not specified) the sample colour changes to yellow. The absorbance was read immediately at 450 nm using an absorbance microplate reader (SpectraMax 340PC, Molecular Devices, Sunnyvale, CA, USA). Standard curve dilutions of the kit supplied p24 control were prepared in complete DMEM and samples measured in one or two dilutions (in complete DMEM) depending on levels of infectious titer (1:4 up to 1:5000).

2.4.4. Vector Genome Titration by Quantitative Reverse Transcription Polymerase Chain Reaction (Q-RT-PCR)

Q-RT-PCR was used to quantify copies of the vector genome RNA in LV producer cells and in LV harvests made in experiments studying effects of producer cell shRNA treatment on vector production (chapter 5).

2.4.4.1. Vector Genome RNA Extraction for Q-RT-PCR in LV Producer Cells

For Q-RT-PCR of vector genome RNA from producer cells, cell pellets of 5 to 10 x10^6 cells were trypsinised (Sigma-Aldrich) from wells of 6-well plates and pelleted by centrifugation at 300 x g at 4°C for 3 min. The RNAeasy mini kit (Qiagen, Limburg, Netherlands) was used to extract total cellular RNA by adding 350 µl lysis buffer RLT in the presence of ß-mercaptoethanol (Sigma-Aldrich) to each cell pellet. Cells were disrupted by freezing at -80 °C and subsequent thawing of the sample on ice. One volume of 70% ethanol (HAYMAN Speciality Products, Witham, UK) was added to each tube followed by RNA purification using Qiagen spin tubes that contain a silica-based membrane. The concentration of eluted RNA was determined by measuring the sample's absorbance at 260 nm using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific) and RNA concentration was normalised to 100 ng/µl with nuclease free water.

2.4.4.2. Vector Genome RNA Extraction for Q-RT-PCR in LV harvests

For Q-RT-PCR from vector harvests (producer cell supernatants) the QiaAmp Viral RNA kit (Qiagen) was used to extract viral RNA. Frozen samples that had been stored at -80°C were thawed on ice and 140 µl of sample was mixed with 560 µl AVL buffer, supplied by the kit, and incubated for 10 minutes at room temperature. This ensured lysis of vector particles and inhibition of RNases. Carrier RNA (1.1 µg/reaction; supplied by the kit) had been added to the AVL buffer to protect the relatively low amounts of viral RNA from degradation by residual RNase activity and to increase its binding to and recovery from the silica-based spin column membrane. A volume of 560 µl absolute ethanol (HAYMAN Speciality Products) was added to each sample and viral RNA purified using the Qiagen spin columns.

An equal volume 10 µl of a total of 60 µl eluted viral RNA was used for gDNA elimination and RT reaction as described below. Negative control was cell supernatant from 293T cells that had not been transiently transfected with vector DNA.

2.4.4.3. Principle of Q-RT-PCR

Quantitative reverse transcriptase polymerase chain reaction (Q-RT-PCR) was applied to quantify RNA copies in vector harvests or cell lysates. A two-step protocol was used: 1) reverse transcription of sample RNA into cDNA and 2) Q-PCR of cDNA. The cDNA target sequence, here a 273 bp product, is amplified in a PCR in the presence of a DNA dye (SYBR® Green) binding double stranded DNA. During the PCR reaction SYBR® Green binds the amplified double-stranded DNA molecules, resulting in the emission of fluorescence. A passive reference dye, ROX™, is present compensating for non-PCR–related variations in fluorescence detection caused by slight variations between wells due to reaction volume variations or the position of the well. The reference dye provides a stable baseline to which PCR-related fluorescent signals are normalised. It is necessary when using any light cyclers from Applied Biosystems, such as the 7500 Fast Real-Time PCR System that was used in these experiments. The intensity of the SYBR® Green fluorescent signal is directly related to the logarithmic amount of amplified cDNA target sequence and increases in a linear fashion compared to the number of PCR cycles. The number of the PCR cycle after which the fluorescent signal rises above the baseline fluorescent signal and starts to increase exponentially is the threshold cycle (Ct). The Ct of the sample Q-PCR is then compared to the Ct of the standard reaction Q-PCR in which the amount of DNA the copy number at the beginning of the PCR is known. PCR standard was the pHV DNA plasmid diluted in nuclease free water to 10^6, 10^5, 10^4, 10^3 and 10^2 copies per µl. The calculation for pHV plasmid DNA standard is shown in Table 4. In knock-down experiments STAR-RDpro-pHV as well as 293T transient producer cells package LVs carrying the vector genome HV, hence for measuring the number of vector genome RNA the same primer pair can be used in the Q-RT-PCR reaction. The primer SFFV Fw and SFFV Rev (Table 3) bind to the region upstream of the SFFV promoter driving GFP expression and within the SFFV promoter, respectively. To quantify vector genome RNA copies in cell lysates results were normalised to human ß-actin mRNA copies (primer: HB actin Fw and HB actin Rev, designed by Sean Knight,

Table 3). The standard plasmid was made by Sean Knight who cloned the PCR product of HB actin Fw and Rev into pGEM T- easy (Clontech, Mountain View, CA, USA).

2.4.4.4. Reverse Transcription and Q-PCR of Vector Harvests and Cell Lysates of GIPZ Producer Cells

For reverse transcription of viral RNA into cDNA with a recombinant reverse transcriptase the QuantiTect Reverse Transcription Kit (Qiagen, Limburg, Netherlands) was used. Genomic DNA (gDNA) that co-purifies during RNA extraction was eliminated by adding gDNA Wipeout Buffer to each sample followed by incubation at 42°C for 2 min. For reverse transcription of RNA extracted from producer cells 800 ng of total RNA were mixed with the reverse transriptase, oligo-dT and random primer, Quantiscript RT Buffer containing dNTPs and an RNase inhibitor. The samples were then incubated at 42 °C for 15 minutes followed by enzyme inactivation at 95°C for 3 min. For reverse transcription of viral RNA extracted from vector harvest 10 µl of viral RNA (equivalent to viral RNA extracted from 23.3 µl vector harvests) were used. The reverse transcriptase has RNA-dependent DNA polymerase activity synthesising cDNA as well as RNase H activity degrading RNA in RNA:DNA hybrids. To confirm complete gDNA wipe-out for each sample a 'no RT' control reaction was set up and analysed by Q-PCR. This control reaction did contain all components apart from the reverse transcriptase and was used to confirm the complete elimination of gDNA. Wipe-out of genomic DNA was necessary to avoid overestimation of RNA contents caused by amplification of gDNA in the PCR of vector samples.
Q-PCR was performed using QuantiTect® SYBR® Green PCR kit (Qiagen, Limburg, Netherlands). For Q-PCR of cDNA 2 µl of template cDNA prepared from producer cells and vector harvests (equivalent to viral RNA extracted from 2.3 µl vector harvests) were mixed with relevant forward and reverse primer and the QuantiTect SYBR Green master mix (Qiagen, Limburg, Netherlands) (Table 5). The master mix contains a HotStarTaq DNA Polymerase, dNTP mix as well as the fluorescent dyes SYBR® Green I and ROX™. The HotStarTaq DNA polymerase is

inactive at room temperature preventing the formation of misprimed products and primer–dimers during reaction set up and in the first denaturation step. Reactions were performed in a 96 well Q-PCR plate (Thermo Fisher Scientific) sealed with optically clear adhesion film. The set-up of PCR reactions is shown in Table 5. Outlined in Table 6 is the PCR method used. In negative control reactions addition of the polymerase was omitted and set up in duplicate wells ('No RT' controls). All samples were analysed in triplicate wells. An example standard curve of Q-PCR of β-actin in cell lysates after reverse transcription is shown in Figure 5. For data analysis of Q-PCR results in cell lysates the comparative C_T method also referred to as the $2^{-\Delta\Delta CT}$ method was used, with $2^{-\Delta\Delta CT}$ = [(C_T gene of interest -C_T internal control) sample A - (C_T gene of interest -C_T internal control) sample B], in which sample A is SFFV and sample B is beta-albumin. Using this comparative analysis the comparison of C_T values of reverse transcribed viral RNA to C_T values of the DNA plasmid standard curve can be avoided. The comparative C_T method assumes that amplification efficiencies of SFFV and beta-actin Q-PCRs are equal or at least differ by less than 10%; however the amplification efficiencies were not directly compared in these experiments.

2.4.5. Lentiviral Vector Protein Detection using Silver Staining and Western Blotting

Viral proteins were semi-quantified in several samples types. To analyse viral proteins in crude vector samples 2 µl were loaded per gel lane. For analysis of SEC fractions 40 µl of one 4 ml fraction were used. For Western blotting of pooled SEC fractions, fractions were concentrated by centrifugation filtration using Amicon Centrifugal Filter Unit (Millipore, Billerica, MA, USA) at 4000 xg until sample volume was reduced to 40 µl (100 fold concentration). To detect viral proteins in vector preparations for LC-MS/MS analysis 50 ng, 100 ng or 250 ng of the purified, desalted and concentrated samples were analysed. Western blotting was also used for knock-down of host cell and viral protein detection. For establishment of knock-down efficiency by

shRNAs targeting selected cellular proteins 10 µg to 30 µg of total protein of LV producer cell lysates were loaded per gel. 45 µl of corresponding vector harvests were analysed to evaluate p24 protein levels.

Samples were subjected to SDS polyacrylamide electrophoresis (SDS-PAGE) using 10% acrylamide gels (Table 7 and 8). Proteins were separated on 1mm thick gels (150 V, 1.5 hours) and visualised by silver staining using the Silver Stain PlusOne kit (GE Healthcare, Little Chalfont, UK) or SilverXpress kit (Invitrogen), according to the provided protocol. For detection by Western blotting proteins were separated on 1.5 mm thick gels (150 V, approximately 2 hours). For Western blotting, proteins were transferred (200 mA, 2.5 hours) from the gel to a Hybond enhanced chemiluminescence (ECL) nitrocellulose membrane (GE Healthcare, Little Chalfont, UK). Membranes were blocked by incubation for 1 hour in PBS with 0.1% Tween-20 (Sigma-Aldrich) and 5% milk powder (Marvel, Original Dried Skimmed Milk Powder) (PBS/0.1%T/5%milk) on a horizontal shaker, followed by binding of specific proteins to the primary antibody. Membranes were incubated overnight sealed in transparent film with the relevant detecting antibody diluted in PBS/0.1%T/0.5%milk. Antibodies used were: anti-p24 (1:100, mouse monoclonal, ARP 365, CFAR, NIBSC, UK; binds sequence NPPIPVGEIY in p24 of HIV-1 Gag), anti-VSV-G (1:2000, mouse monoclonal, Sigma-Aldrich), anti-RDpro pg70 (1:2000, goat polyclonal, Quality Biotech Inc., Camden, NJ, USA), anti-AHANAK (1:500, mouse-monoclonal, Abcam, Cambridge, UK), anti-TSG101 (1:500, rabbit-polyclonal, Sigma-Aldrich), anti-ALIX (1:1000, rabbit polyclonal), anti-EEF1A (1:500, mouse-monoclonal), anti-ENO1 (1:500, mouse-monoclonal), anti-MARCKSL1 (1:1000, rabbit-polyclonal) and anti-GAPDH (1:5000, mouse-monoclonal, all by Millipore, Billerica, MA, USA). Proteins were transferred on to a nitrocellulose membrane (Hybond enhanced chemiluminescence (ECL) nitrocellulose membrane (GE Healthcare)) for 1 hour (or 2.5 hours for the transfer of AHNAK) at 200 mA. Membranes were washed for 3x 10 minutes in PBS/T. Either horseradish-Peroxidase (HRP) conjugated secondary anti-mouse or anti-rabbit antibodies were incubated with the membranes diluted 1:2000 in PBS/T/0.5%milk for 1 hour (Sigma-Aldrich). Followed by washing

the membrane for 3x 10 minutes in PBS/T, protein signals were developed using ECL Western Blotting Detection Reagents and exposed to ECL Hyperfilm (both by GE Healthcare, Little Chalfont, UK). For densitometry readings of Western blot signals the software Image J (version 1.46r) was used.

2.5. Knock-Down of Cellular Protein Expression

2.5.1. GIPZ shRNA-Lentiviral Vector Preparation

Lentiviral vector particles carrying the pGIPZ vector genome (LVs-GIPZ) were prepared by co-transfection of packaging plasmid p8.91, VSV-G envelope expression plasmid pMDG and the SIN-lentiviral vector genome plasmid, pGIPZ, carrying a shRNA encoding a part of the antisense sequence for the mRNA targeted protein. PGIPZ plasmids were purchased from Thermo Fisher Scientific and provided as *E.coli* bacteria stab cultures. Target sequences of all shRNAs used are listed in Table 9. The plasmid contains the LV genome elements 5' and 3' LTR, central polypurine tract (cPPT), the packaging sequence Ψ and a WPRE. The shRNA is based on the endogenous miRNA 30 (mir-30). ShRNA expression is marked by the reporter gene turbo green fluorescent protein (tGFP) driven by the cytomegalovirus (CMV) promoter upstream of an internal ribosome entry site (IRES) driving the expression of puromycin. PGIPZ plasmid preparation is described in section 2.2.3. Each stab culture was streaked out on an LB agar plate containing 100 µg/ml ampicillin and incubated overnight at 37°C. The next day 5 ml of LB (with 100 µg/ml ampicillin) were inoculated with a single colony picked of the agar plate and incubated for approximately 8 hours at 37°C on a shaker at 300 rpm. Plasmid DNA was extracted from each culture using a DNA plasmid extraction kit (Qiagen, Limburg, Netherlands).

293T cells were plated in 10 cm dishes at a density of $3x10^6$ cells per dish in 10 ml complete DMEM., 24 hours after plating, medium was changed to 8ml complete DMEM and cells were

transiently transfected with 1µg of p8.91, 1 µg of pMDG and 1.5 µg of pGIPZ as well as 36 µl Fugene transfection reagent in 800 µl Opti-MEM® Reduced Serum Medium, GlutaMAX™ per 10 cm dish. Twenty-four hours after transfection the medium was changed to 8 ml complete DMEM. The first time cell supernatant was collected at 48 hours post transfection and stored at 4°C overnight followed by the second collection at 72 hours. First and second collection of supernatants were pooled, filtered (0.2 µm cellulose acetate, Schleicher & Schwell Micro Science) and concentrated 40 fold by ultracentrifugation at 48 000 x g for 2 hours (Beckman Coulter, L-90K, SW28) at 4°C. Aliquots were stored at -80°C. Titration of GIPZ-LV preparations was carried out as described in section 2.3.2.

2.5.2. Transduction of LV Producer Cells with GIPZ-LVs to Deliver shRNA

STAR cells or 293T cells were plated in 6 well plates at a density of $6x10^5$ cells per well and transduced 24 hours later with pGIPZ-LVs at MOI 10 in complete DMEM medium with 8 µg/ml polybrene. To identify the shRNA with highest knock-down efficiency 293T cells were transduced with one of the shRNAs in Table 9. At least two shRNAs were analysed per target protein. Transduced cells were selected in puromycin containing medium (10 µg/ml) starting twenty-four hours after transduction and maintained from then onwards in puromycin conditioned medium.

2.5.3. Production of LVs by STAR-GIPZ and 293T-GIPZ Cells

After a minimum of one week in selection medium STAR-GIPZ and 293T-GIPZ cells were seeded at a density of $6x10^5$ cells per well of a 6-well plate. For the transfection of 6 wells of 293T-GIPZ cells 24 hours later 1 µg pMDG, 1 µg p8.91 and 1.5 µg of pHV were mixed with 36 µl Fugene (Roche, Burchess, Hill, UK) and 800 µl Opti-MEM® Reduced Serum Medium, GlutaMAX™ (Invitrogen) after medium change to 1.33 ml of fresh DMEM per well. The next

day the medium was changed to 2 ml fresh DMEM and 48 hours and 72 hours post transfection 2 ml of cell supernatant were collected. After filtration through 0.45 µm cellulose acetate syringe filters (Schleicher & Schwell Micro Science) vector harvests were frozen and stored at -80°C for later analysis. For vector harvests from STAR-GIPZ producers the medium was changed to 2 ml of fresh medium 48 hours after plating and 24 and 48 hours later (day 4 and 5 after plating) supernatant collected, filtered and stored as described above. At the time of the second vector harvest (day 5 after plating), producer cells were lysed and target host protein expression was analysed by Western blotting.

2.5.4. Preparation of Cell Lysates

Cell lysates for knock-down analysis by Western blotting was prepared from STAR and 293T cells as well as STAR-GIPZ and 293T-GIPZ cells of knock-down analysis experiments (section 2.5.3.; experiments of chapter 4). A confluent well of a 6-well plate of cells was washed with 500 µl cold PBS. Cells were lysed by adding 150 µl of cold cell lysis buffer. On ice the plate was shaken for 10 minutes before lysates were collected and centrifuged at 13 000 xg at 4°C for 20 min. Supernatants were transferred to pre-chilled 1.5ml tubes and lysates stored at -80°C. Alternatively protein concentration was determined immediately using the bicinchoninic acid (BCA) assay (Thermo Fisher Scientific).

2.5.5. Protein Concentration Measurement using BCA assay

Total protein concentration was measured using the bicinchoninic acid (BCA) assay (Thermo Fisher Scientific) in the microplate format. Following the manufacturer's protocol standard dilutions of 2000 to 25 µg/ml of bovine serum albumin (BSA; Thermo Fisher Scientific) were made in protein lysis buffer and 20 µl of each standard as well as 20 µl of cell lysate sample (diluted 1:5 in protein lysis buffer) used in the assay. The absorbance was measured at 562 nm using a microplate reader (SpectraMax 340PC; Molecular Devices, Sunnyvale, CA, USA).

Samples were diluted in protein lysis buffer to a concentration of 1 µg/µl and SDS-PAGE sample buffer added before storage at -20°C for further analysis by Western blotting.

2.5.6. Functional Category and Canonical Pathway Analysis of LC-MS/MS Data

IPA (Ingenuity® Systems [www.ingenuity.com], Mountain View, CA, USA) was used to determine the most significant molecular and cellular functions and cellular pathways in the LC-MS/MS datasets. Datasets in the form of Excel spreadsheet lists of UniProt identifiers were uploaded into the application. For core analysis of a dataset default settings were applied and only direct relationships between proteins that have been experimentally confirmed considered.

2.6. Materials and Methods - Appendix

Table 1: Vector Samples for LC-MS/MS Analysis

Lentiviral Vector Sample	Components
VSV-G-GFP	transiently produced, VSV-G Env, Gag/Pol, vector genome (pRRLSIN.cPPT.PGK-GFP.WPRE; transgene: GFP) and Rev
VSV-G-Empty	transiently produced, VSV-G Env, Gag/Pol protein and Rev
Gag/Pol-GFP	transiently produced, Gag/Pol, vector genome (pRRLSIN.cPPT.PGK-GFP.WPRE; transgene: GFP) and Rev
VSV-G only	transiently produced, VSV-G Env, vector genome (pRRLSIN.cPPT.PGK-GFP.WPRE; transgene: GFP); control sample
RDpro-GFP	stably produced by STAR-RDpro-pHV expressing RDpro envelope protein, Gag/Pol, vector genome (pHV, transgene: GFP), Tat and Rev
RDpro-Empty	stably produced by STAR-RDpro, expressing RDpro envelope protein, Gag/Pol, Tat and Rev, no vector genome

Table 2: DNA plasmids used in Transient Transfections

Plasmid	Source	comments
p8.91	(Zufferey et al., 1997) Plasmid Factory (Bielefeld, Germany)	Expression plasmids for CMV driven Gag-Pol; used in transient transfection experiments chapter 5 to prepare GIPZ-LVs
pMDG	(Naldini et al., 1996) Plasmid Factory	Expression plasmids for CMV driven VSV-G; used in transient transfection experiments chapter 5 to prepare GIPZ-LVs
RD-pro	Y. Ikeda (Ikeda et al., 2003, Cosset et al., 1995)	Expression plasmids for MLV-LTR driven RDpro
ß-actin	Sean Knight	Expression plasmid for cDNA of exon 4 of human ß-actin driven by T7 promoter
pHV	(Ikeda et al., 2003, Kotsopoulou et al., 2000)	LV genome; 5'LTR: CMV-R-U5; 3'LTR: U3-R-U5; SFFV driving GFP
pGIPZ	Thermo Fisher Scientific	SIN-LV genome, shRNA gene
pMD2.G	Addgene (Cambridge, MA, USA) # 12259 (Trono lab)	Expression plasmid for CMV driving VSV-G; used in transient transfection experiments chapter 3
pRSV-Rev	Addgene (Cambridge, MA, USA) # 12253 (Dull et al., 1998)	Expression plasmid for RSV driving Rev
pMDLg/pRRE	Addgene (Cambridge, MA, USA) #12251 (Dull et al., 1998)	Expression plasmids for CMV driven Gag-Pol; used in transient transfection experiments chapter 3
pRRLSIN.cPPT.PGK-GFP.WPRE	Addgene (Cambridge, MA, USA) # 12252 (Follenzi et al., 2000)	SIN-LV genome; 5'LTR: CMV-R-U5, 3'LTR: ΔU3-R-U5; PGK driving GFP

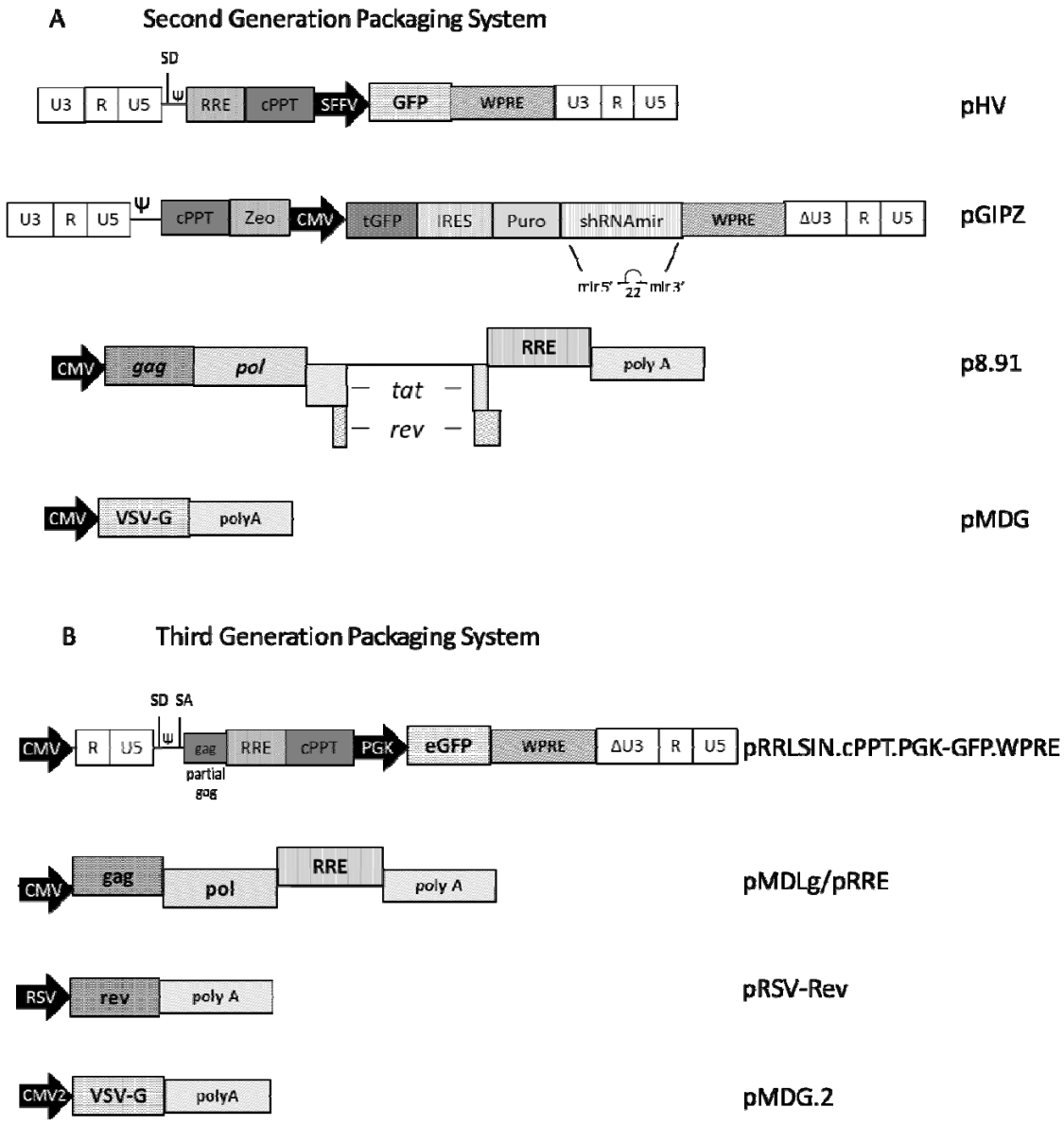


Figure 5: Map of DNA plasmids after integration into producer cells for transient vector production showing LV relevant elements only: A) Plasmids of the Second Generation Packaging System, pGIPZ, p8.91 and pMDG were used for preparation of GIPZ-LVs (section 2.5.1). Plasmids pHV, p8.91 and pMDG were used for transient transfection of 293T-GIPZ cells (section 2.5.3). B) Plasmids of the Third Generation Packaging System, pRRLSIN.cPPT.PGK-GFP.WPRE, pMDLg/RRE, pRSV-Rev and pMDG.2 were used for transient transfection of 293T cells (section 2.3.1).

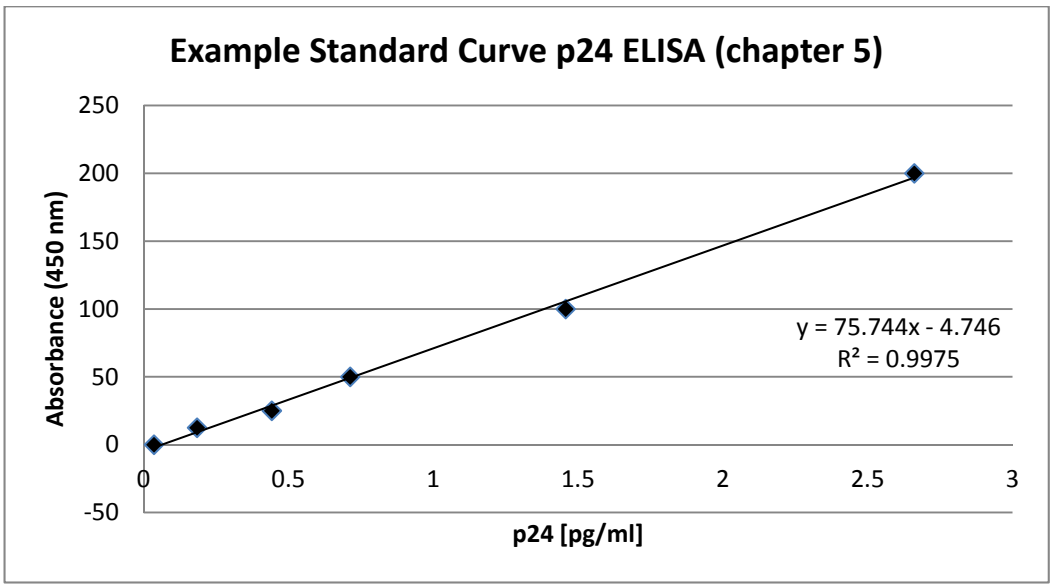


Figure 6: Example of standard curve for p24 ELISA (using Lenti-X™ p24 Rapid Titer Kit chapter 5).

Table 3: Primer for Q-PCR of Vector RNA Genome.

Target mRNA	Primer Name	Primer Sequence
SFFV	SFFV Fw	CGATAAGCTTGATATCGAATTCCT
	SFFV Rev	TGCGGTGACCATCTGTTC
1. Human ß-actin	HB actin Fw	TGGACTTCGAGCAAGAGATG
	HB actin Rev	GAAGGAAGGCTGGAAGAGTG

Q-PCR primer (purchased from Invitrogen): SFFV primer bind to the region upstream of the SFFV promoter driving GFP expression and within the SFFV promoter. Vector genome RNA copy numbers were normalised to ß-actin copy numbers in cell lysates.

Table 4: Calculation of pHV Plasmid DNA Standard Molecular Weight Used in Q-PCR of Vector RNA Genome.

Standard	Plasmid Size (bp)	Molecular Weight (Da)	Mass (g) / copy	Weight of 10^{10} copies (ng)
pHV	7918	5225880	8.67×10^{-18}	86.78
ß-actin	3146	2076360	3.45×10^{-18}	34.48

The size of the plasmid was used to calculate the molecular weight of one vector genome copy (with 1 bp = 660 Da and 1Da = 1.66×10^{-24} g) from which the mass of 10^{10} in nanograms (ng) was calculated and used to make a 10^{10} plasmid/µl aliquot in Buffer EB (Qiagen), this was serially diluted to give aliquots of 10^6, 10^5, 10^4, 10^3 and 10^2 plasmids/µl.

Table 5: PCR Reaction Set-Up.

Component	Concentration	Volume per reaction (µl) - Standards	Volume per reaction (µl) - Samples
Primer F	20 µM (final 0.4)	0.5	0.5
Primer RC	20 µM (final 0.4)	0.5	0.5
QuantiTect SYBR Green master mix	na	12.5	12.5
RNase-Free Water	na	6.5	9.5
Template cDNA (added in to each reaction individually)	na	5	2
Total Reaction Volume	na	25	25

Standard and sample reactions differ in volume of template added; 5 µl and 2 µl were added, respectively.

Table 6: Q-PCR Program

Step	Time	Temperature (°C)
PCR initial heat activation	15 min	95
3-step cycling (40 cycles)		
Denaturation	15 sec	94
Annealing	30 sec	55
Extension	30 sec	72
Melting curve		

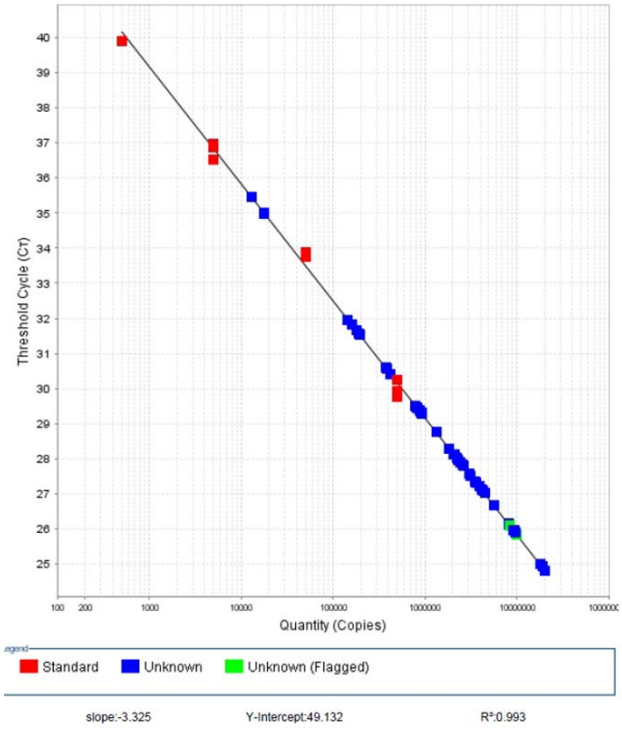


Figure 7: Example standard curve of Q-PCR of β-actin after reverse transcription of cell lysate RNA, see section 2.4.4.4. Efficiency = $10^{(-1/slope)}$-1 = $10^{(-1/-3.325)}$-1 = 99.8%.

Table 7: Composition of Acrylamide Gels used in Western Blotting.

Component	Separating Gel (10%)	Stacking Gel (4%)
dH2O	7.94 ml	6.8 ml

Tri-HCL	5 ml of 1.5M (pH 8.8)	1.25 ml of 0.5M (pH 6.8)
30 % Acrilamide-Bisacrylamide (Bio-Rad, Hercules, CA, USA)	6.6 ml	1.7 ml
20% SDS (media facility, NIBSC)	100 µl	50 µl
APS	200 µl	100 µl
TEMED	20 µl	10 µl

All reagents were purchased from Sigma-Aldrich unless stated otherwise.

Table 8: Composition of Buffers Used in Western Blotting.

Buffer	**Composition**
Phosphate Buffered Saline (PBS)	10 mM sodium phosphate dibasic, 156 mM sodium chloride, 2mM potassium phosphate monobasic, pH 7.2 to 7.4
Cell lysis buffer	1% Triton X-100 in PBS (Sigma-Aldrich) containing protease inhibitors: cOmplete(R)^{TM} EDTA-free, broad spectrum inhibitors of serine and cysteine proteases (Roche, Burchess, Hill, UK)
SDS-PAGE sample buffer	50 mM Tris-HCL, pH 6.8, , 10% glycerol, 2% SDS, 12.5 mM EDTA, 0.02% bromophenol blue, 1% ß-mercaptoethanol
SDS-PAGE running buffer (10x)	250 mM Tris base, 1.92 M glycine, 1 % SDS, pH 8.3
Transfer buffer (10x)	250 mM Tris base, 1.92 M glycine, pH 8.3

All reagents were purchased from Sigma-Aldrich unless stated otherwise.

Table 9: pGIPZ-shRNA Plasmids.

Target protein	**shRNA #**	**pGIPZ ID**	**shRNA target (antisense)**
ALIX	2	V2LHS_64522	TTTCGTTGCAGTAATTCAG

	3	V2LHS_64525	TGATCTGTTAACCAAATCG
	11	V2LHS_64521	TTAGTCAATATAGACTGAG
	18	**V2LHS_64526**	**TAATCTGCAGCCTGATTAG**
AHNAK	1	V2LHS_233334	TAACTGCAGGTGTTTGTTG
	6	**V2LHS_92221**	**TAGATCAGGAGCTCCTACG**
	8	V2LHS_160999	TTATGTCAATTTCAGGGCC
	10	V2LHS_98839	TCCAGTGCTGATGGCTGTG
	13	V2LHS_98840	TTGCATTCCAGTGCTGATG
	38	V3LHS_360956	GAGACAACAACATCAGCCT
	39	V3LHS_360958	AGAAGAGGAGGACAGTCGG
	41	V3LHS_360959	TAAATTGAAATCAACATCA
EEF1A	4	V2LHS_113620	TTGTAGACATCCTGGAGAG
	29	V3LHS_360005	TACAATACCGGTAACAACG
MARCKSL1	27	V3LHS_412143	TGAACTTGAGTAAGACATT
	20	V2LHS_259022	TCTTCTGAGGCTGCACTAG
	26	V3LHS_412142	AAGGGACCATCTTCAACTG
	42	V3LHS_412141	AAGACATTTATAAAAACCT
ENO1	17	V2LHS_113737	TTGTCCCGGAGCTCTAGGG
	37	V3LHS_408476	TTGAGCACAAAACCACCGG
	40	V3LHS_408477	TTTGAGCACAAAACCACCG
TSG101	45	V3LHS_305576	ACTTCTTGATCTAAACGGG
	46	**V2LHS_56429**	**TGCAATAACTTATTCTGGG**
Non-silencing control			**ATCTCGCTTGGGCGAGAGTAAG**

Selected host proteins were targeted with at least two shRNAs to determine which shRNA achieves the highest protein knock-down. PGIPZs were purchased from Thermo Fisher Scientific using the pGIPZ ID. Highlighted in green are shRNA sequences that were used in experiments analysing host protein knock-down effects on vector production.

3. Lentiviral Vector Production and Purification by Size Exclusion Chromatography

3.1. Introduction

In several clinical trials lentiviral vectors (LVs) have been shown to successfully treat several diseases, including rare inherited diseases for which there are currently no alternative long-term therapies available other than allogeneic hematopoietic stem cell transplantation (HSCT). Clinical trials for the treatment of monogenic disorders such as beta-thalassaemia (Cavazzana-Calvo et al., 2010), Wiscott-Aldrich syndrome (Biffi et al., 2013) and Metachromatic Leukodystrophy (MLD) (Biffi et al., 2013) successfully used *ex vivo* transduction of haematopoietic stem cells (HSCs) with LVs delivering the correct form of disease related genes. LVs also have the potential to be applied as vaccines as they can efficiently transduce dendritic cells (Lopes et al., 2011, Yang et al., 2011, Rodriguez et al., 2012). Consequently, LVs will not only be used increasingly in later stages of clinical trials with larger cohorts of patients, they are also being developed for applications targeting larger groups of individuals.

With the increasing need for large amounts of LVs, their design and method of production are constantly under development. Vector production in current clinical trials and in most laboratory based applications has been based on transient transfection of 293T cells. High titer vectors can be produced; however this lasts only for a few days of production and can vary between batches. Manufacturing vectors by good manufacturing practice (GMP) with this method is expensive due to the requirement of high amount of plasmid DNA, transfection reagents and labour time and difficult to scale up (Merten et al., 2011).

A packaging cell line stably expressing LVs would circumvent these problems. The development of such stable producer cells is based on the stable expression of all packaging components followed by introduction of the transfer vector by either transfection or

transduction. Several packaging cell lines have been developed but due to toxicity of vector components, resulting from overexpressed viral proteins such as Rev and VSV-G in producer cells, tetracycline-dependent inducible systems have been used. Vector production, specifically viral gene expression is regulated by either withdrawal or addition of tetracycline (Broussau et al., 2008, Stewart et al., 2009, Stewart et al., 2011). However addition or removal of induction reagents can complicate downstream purification processes. Furthermore difficulties in the stable expression of HIV-1 *gag-pol* impaired the development of a high titer packaging cell line.

The producer cell line STAR stably expresses high levels of a codon optimised HIV-1 *gag-pol* introduced by transduction with murine leukaemia virus (MLV) derived vectors as well as *rev* and *tat* (Ikeda et al., 2003). The commonly used VSV-G Env has a wide cell tropism and good stability but is known to be cytotoxic when stably expressed in producer cells (Burns et al., 1993). Alternatively the γ-retroviral envelope RDpro, derived from RD114 Env, was used in STAR cells as it can efficiently transduce HSC and is not cytotoxic (Relander et al., 2005, Bell et al., 2010). Further improvements of the STAR cell line need to be undertaken since produced LVs are less infectious compared to VSV-G pseudotyped particles with transduction titers about ten-fold below transiently produced (Ikeda et al., 2003, Strang et al., 2004, Bell et al., 2010).

In order to analyse the protein content of LVs they need to be separated from non-associated proteins. Several methods for viral vectors downstream processing have been described that involve vector concentration and purification from producer cell supernatant. Microfiltration is often used to clarify vector containing supernatant and serves as the initial step of vector purification immediately after vector harvest. Producer cells and cell debris can be removed from the cell culture supernatant by filtration through membranes with a pore size of 0.2 µm and 0.45 µm (Segura et al., 2005, Segura et al., 2006b). Ultrafiltration is a very gentle method to concentrate large amounts of vector containing supernatant and employs a membrane such as polyethersulfone with a high molecular weight cut-off (for example 300 kDa). This can be

followed by diafiltration to exchange the buffer for subsequent chromatography (Segura et al. 2005). Parameters such as pressure, stirring rate and process times need be kept low in order to have satisfying recovery rates (Cruz et al., 2000). Tangential flow modes, using hollow fibres, are easier to scale up as build-up of solids on the membrane is low due to tangential flow of the retentate to the membrane, allowing a higher flow rate (Geraerts et al., 2005). Alternatively ultracentrifugation concentrates vectors from supernatant by separation of vectors using centrifugal force allowing the removal of cell culture media (Burns et al., 1993). However proteins with similar or greater density to viral particles co-concentrate during the process. Vectors purified by sucrose-density ultracentrifugation have been shown to be contaminated with microvesicles that have a similar density to vectors (Bess et al., 1997). Furthermore this method is also cumbersome, difficult to scale up and reduces functional particles yields significantly (Baekelandt et al., 2003). Using rate zonal ultracentrifugation, vectors are separated by size and density from contaminants including defective vectors and cell membrane vesicles, however this method achieved a relatively low recovery of infectious particles (37%) (Segura et al., 2006a). Furthermore, since the size of microvesicles is thought to range between 50 to 500 nm, vesicles of similar size to vectors would not be removed by this method.

These initial steps in vector downstream processing are primarily focused on removal of cells, cell debris and the reduction of harvest volume. They are followed by the actual purification step removing most impurities in cell supernatants from vectors. For the elimination of plasmid DNA and genomic DNA from host cells current methods used for large scale LV clinical grade purification include benzonase (endonuclease) treatment followed by membrane based anion-exchange chromatography (AEXc) (Schweizer and Merten, 2010, Aiuti et al., 2013, Biffi et al., 2013). Anion-exchange chromatography is the absorption of the negatively charged vectors to the positively charged chromatographic supports such as diethylaminoethanol (DEAE) anion-exchangers (Merten et al., 2011, Aiuti et al., 2013, Biffi et al., 2013). Unconcentrated vector supernatant can be loaded directly onto the column. The volume of

applied supernatant is limited by the protein binding capacity of the anion-exchange support. Based on quantitative PCR data of the vector genome of eluted particles recovery rates were low, ranging from 22 % to 68% (Scherr et al. 2002). Instead of agarose or resin beads the column matrix can consist of a membrane. Mustang Q anion-exchange cartridges contain an anion-exchange support on a membrane with a large pore sizes (0.8 µm) and filtered supernatants can be loaded. For laboratory scale vector processing of up to 1500 ml vector harvests can be processed in Mustang Q ion exchange filter capsules with a capacity of 900 ml (Slepushkin et al., 2003). Capsules with a capacity of 5 litres are now available (GE Healthcare, Little Chalfont, UK). Rodrigues et al. (2008) showed that viral particles can be separated from envelope protein-free vectors and soluble envelope proteins due to their lower binding strengths at physiological pH. Vectors pseudotyped with envelopes other than VSV-G can also be purified such as gibbon ape leukaemia GALV-Env (Rodrigues et al., 2008) or RD114 and rabies virus Env glycoproteins (Kutner et al., 2009) making this techniques applicable for a wide range of viral vectors. Other chromatography methods have been used to purify clarified and concentrated vector preparations. Affinity chromatography using heparin affinity matrices allows vector particle elution at moderate salt concentrations compared to AExc (350 mM and 900 mM NaCl, respectively). To prepare concentrated vectors for affinity chromatography an extra step of ultrafiltration into heparin absorption buffer is required. Heparin affinity chromatography purified vectors are only partly purified and polishing by size exclusion chromatography is required (Segura et al., 2005).

Size exclusion chromatography (SEC) separates proteins according to the hydrodynamic volume of their molecules. Larger molecules migrate faster through the matrix, which consists of porous, non-absorbing material, such as sepharose. This method was used during processing of a HIV-1 based VSV-G pseudotyped lentiviral vector in a phase I clinical trial (Levine et al., 2006). SEC has a dilution effect that requires the feed and also the eluate to be concentrated after chromatography. Infectious particle recovery is high (80%) (Slepushkin et al. 2003), though this can be significantly reduced during subsequent concentration steps.

Other factors, such as low flow-rate, which require long processing times, make this method difficult to use in bulk production of LVs and is more suitable in small scale experiments or as a final polishing step (Segura et al. 2005). Vector sample production protocols currently used for clinical applications use SEC for polishing to remove salts and small sized-contaminants as well as a reformulation step, in order to change medium (Aiuti et al., 2013, Biffi et al., 2013). According to suppliers specification macromolecules in the range of $4x10^4$ to $2x10^7$ relative molecular mass (M_r) can be separated using Sephacryl 500-HR media. It is estimated that immature and mature HIV-1 virions have an average mass of 277 MDa (or $2.7x10^8$ M_r) (Carlson et al., 2008), hence vector particles are above the exclusion limit of $2x10^7$ M_r and expected to elute in the void peak.

For downstream processing of LVs produced in this study a combination of microfiltration using 0.45 µm cellulose acetate filters to separate vectors from cells and cell debris, followed by concentration by ultracentrifugation and SEC purification was employed. Purified vectors were then desalted by dialysis and further concentrated by lyophilisation.

3.2. Aim

The aim of this chapter was the production of LVs and their purification using size exclusion chromatography. Purified samples are characterised regarding their viral protein content. Two different vector production methods are directly compared: the transient transfection method using calcium phosphate precipitation (Ca_2PO_4) adopted for the production of LV in clinical trials and LV production from the stable producer cell line STAR (Ikeda et al., 2003). Preparation of purified vectors for subsequent protein content analysis by liquid-chromatography mass spectrometry (LC-MS/MS) is also presented.

3.3. Results

Six different vector samples were prepared (see chapter 2, section 2.3.1) to compare host cell proteins in transiently and stably produced vectors (Figure 6). Transiently produced vectors were 1) VSV-G-GFP: by transfecting plasmids coding for VSV-G Env, structural protein Gag/Pol, a vector genome with GFP marker gene and Rev; 2) VSV-G-Empty: by transfecting plasmids coding for VSV-G Env, Gag-Pol and Rev, 3) Gag/Pol-GFP: by transfecting plasmids coding for Gag/Pol, vector genome and Rev. 4) VSV-G only: a sample carrying only the VSV envelope was prepared by transient transfection of 293T cells with the VSV-G expression plasmid and the SIN-LV genome plasmid. Gag/Pol-GFP was used as an envelope free-control sample hypothetically lacking host proteins associated with VSV-G or RDpro envelope protein. The VSV-only control was used to potentially identify proteins that are associated with VSV-G envelope protein. Vectors were stably produced by STAR-RDpro cells continuously producing LVs, by expressing RDpro envelope protein, Gag/Pol, vector genome with marker protein GFP and Rev (=RDpro-GFP, sample 5) (Ikeda et al., 2003) or without a vector genome (RDpro-Empty, sample 6).

A transiently produced vector sample RDpro-GFP, produced by transfection of RDpro Env, structural protein Gag/Pol, a vector genome with GFP marker gene and Rev would have completed the set and potentially allowed to distinguish between host proteins that associate with VSV-G or RDpro Env, but was not included due to time constraints.

Figure 8: Summary of five lentiviral vector samples for the study of associated cellular proteins. Transiently produced samples include two VSV-G-pseudotyped samples (with or without a vector genome) as well as samples of vector particles without an envelope protein (Gag/Pol-GFP) or a sample of harvests from 293T cells transiently transfected only with VSV-G Env and

vector genome plasmid DNA, omitting Gag/Pol DNA encoding structural proteins (VSV-G only). Two stably produced samples were produced by the packaging cell line STAR and include RDpro-pseudotyped vectors with and without a vector genome.

3.3.1 Production of LVs –Differences in Titer of Transiently and Stably Produced Vectors

The transduction titer of five batches of transiently produced VSV-G-GFP and stably produced RDpro-GFP was assessed by transducing 293T cells with concentrated LV preparations followed by measuring GFP positive cells by FACS. The average transduction titer for VSV-G-GFP and RDpro-GFP batches were $34.3\pm9.0x10^6$ TU/ml and $3.4\pm2.9x10^6$ TU/ml, respectively, demonstrating an average about 10 fold higher infectious titer of transiently than stably produced vectors. Vector batches used for LC-MS/MS sample preparations are shown in Table 10.

Table 10: Viral Vector Batches used in Sample Preparation for LC-MS/MS Analysis

	batch #	producer cell line	transfected plasmids	TU/ml (concentration factor)
VSV-G-GFP (1)	55	293T	VSV-G/Gag-Pol/GFP	$24.6x10^6$ (40)
	58	293T	VSV-G/Gag-Pol/GFP	$36.0x10^6$ (40)
	74	293T	VSV-G/Gag-Pol/GFP	$42.3x10^6$ (40)
VSV-G-Empty (2)	56	293T	VSV-G/Gag-Pol	na
	62	293T	VSV-G/Gag-Pol	na
	75	293T	VSV-G/Gag-Pol	na
Gag/Pol-GFP (3)	57	293T	Gag/Pol/GFP	na
	60	293T	Gag/Pol/GFP	na
	63	293T	Gag/Pol/GFP	na
	76	293T	Gag/Pol/GFP	na
VSV-G-only (4)	103	293T	VSV-G/GFP	na
RDpro-GFP (5)	64	STAR_RDpro_pHV	none	$0.9x10^6$ (40)
	67	STAR_RDpro_pHV	none	$2.5x10^6$ (40)
	87	STAR_RDpro_pHV	none	$2.8x10^6$ (40)
	91	STAR_RDpro_pHV	none	$13.3x10^6$ (240)
	93	STAR_RDpro_pHV	none	$2.8x10^6$ (240)
	95	STAR_RDpro_pHV	none	$4.9x10^6$ (240)
RDpro-Empty (6)	65	STAR_RDpro	none	na
	68	STAR_RDpro	none	na
	88	STAR_RDpro	none	na
	92	STAR_RDpro	none	na
	94	STAR_RDpro	none	na
	96	STAR_RDpro	none	na

Transduction titer is stated in transducing units per ml (TU/ml). A batch consists of vector samples harvested from producer cells plated and processed in one experiment. 40 fold concentrated VSV-G pseudotyped transiently produced vector samples had a transduction titer about 10 fold higher compared to 40 fold concentrated vectors produced by stable production.

Concentration factor of stably produced batches was increased to 240 fold to increase transduction titer (section 2.5.).

3.3.2. Size Exclusion Chromatography of Vectors – Elution of LVs in Void

During ultracentrifugation large cellular and subcellular debris are concentrated alongside vector particles and form part of the vector sample that was resuspended in Opti-MEM (hence referred to as crude sample). Size Exclusion Chromatography (SEC) was used to separate vector particles from small proteins. Columns were packed with Sephacryl-500–HR medium with a bed volume of 140 ml and TEN buffer as mobile phase flowing at 0.8 ml/minutes (see section 2.3.2 of Materials and Methods chapter).

For SEC 900 µl of crude vector sample were collected per column. The overlay of representative SEC protein elution profiles for each of the five vector samples is shown in Figure 7. Figure 7 A shows a batch of three transiently produced samples VSV-G-GFP, VSV-G-Empty and Gag/Pol-GFP with a void peak maximum between 0.8 and 1.0 mVolts (0.0008-0.001 absorbance units [AU] at 280 nm), with Gag/Pol-GFP samples having the smallest void peak. Figure 7 B shows the equivalent for a batch of three stably produced samples, RDpro-GFP and RDpro-Empty showing void peak heights at 0.25 mVolts. Void peak heights of stably produced samples ranged from 0.1-0.25 mVolts (0.0001-0.00025 AU at 280 nm), hence they contain a lower amount of total proteins in the void fraction. Using the Sephacryl 500-HR column, the void peak in all samples was well separated from following peaks containing lower molecular weight proteins. Fractions containing the void peak were collected for further analysis and processed for LC-MS/MS. Control samples were 900 µl Opti-MEM that had not been in cell contact and Opti-MEM that was 24 hours incubated with 293T cells (900 µl). SEC chromatograms of these controls (Figure 7 C) showed a void peak of <0.1 mVolts, which is 2.5 fold lower than the void peak height of stably produced samples (Figure 7 B) and up to 10 fold

lower than that of transiently produced samples (Figure 7 A), indicating that the void peak fractions of purified vector samples contain mainly proteins derived from vector production.

Resuspension of vector samples in complete DMEM after ultracentrifugation resulted in high BSA levels in crude vector samples as shown in Figure 7 D. The protein that eluted at ~120 minutes post-injection is believed to be derived from FCS in the complete DMEM as the peak height in samples resuspended in complete DMEM was about 10 fold higher than in vector samples resuspended in serum reduced Opti-MEM (Figure 7 A to C). Therefore reduced serum OptiMEM was used for vector preparation of all samples prepared for LC-MS/MS analysis. The size chromatography profile of 900 µl sample elution buffer TEN confirms all of the peaks represent protein and none of them are background noise (Figure 7 E).

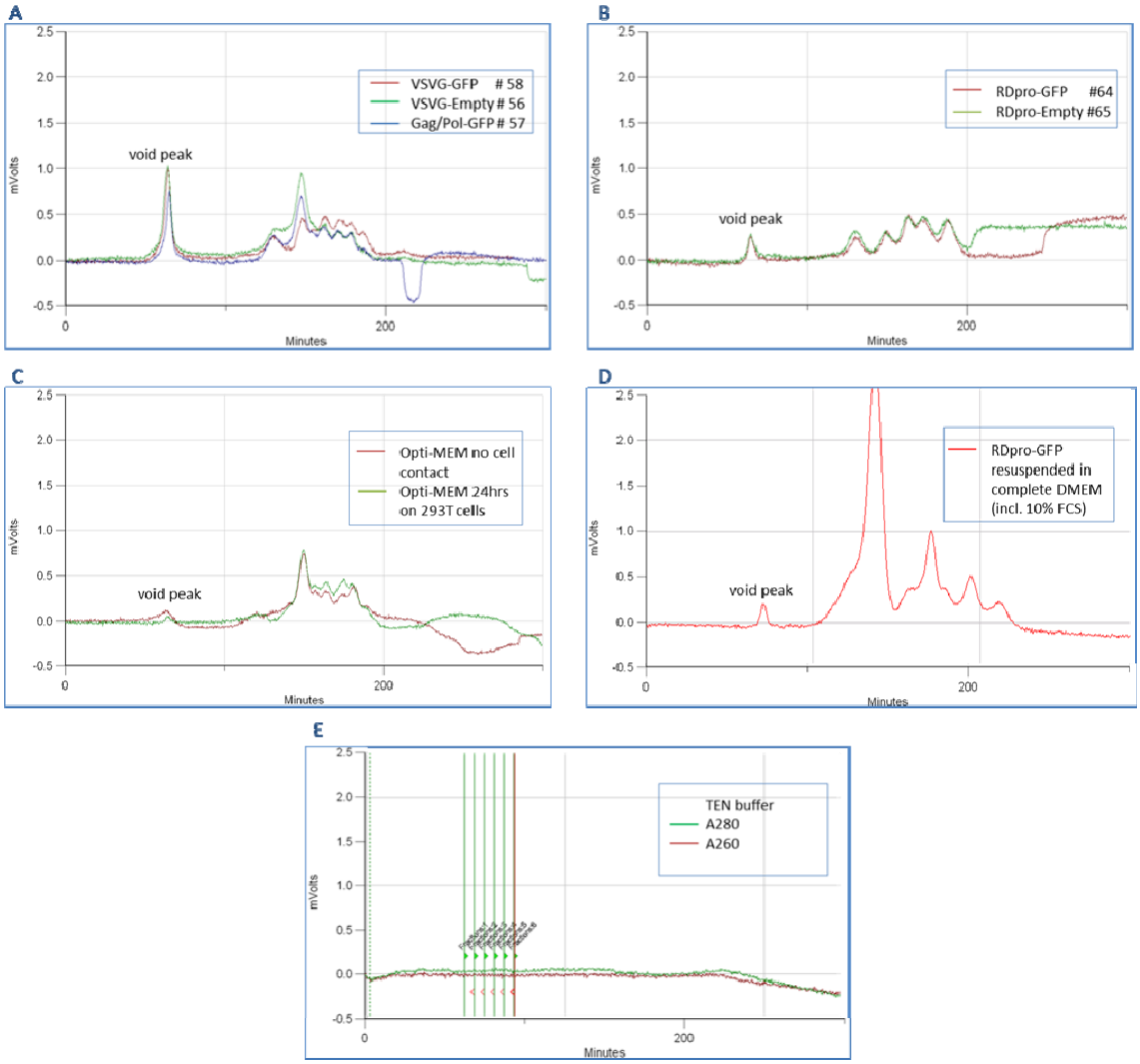


Figure 9: SEC void peaks from cell supernatants collected from vector producing cells contain up to 2.5 to 10 fold more proteins compared to those collected from untransfected 293T cells. SEC chromatograms of vector samples, 900 µl of each concentrated vector sample were applied to the SEC column, representative elution profiles are shown. A) transiently produced vector samples, VSV-G-GFP, VSV-G-Empty and Gag/Pol-GFP, B) stably produced vector samples, RDpro-GFP, RDpro-Empty, batch number (#) of vector preparation is indicated, C) Opti-MEM and Opti-MEM incubated 24hrs on 293T cells, D) RDpro-GFP vector preparations resuspended in complete DMEM (incl. 10 %FCS) after ultracentrifugation instead of serum reduced Opti-MEM, E) TEN buffer. 1 mVolts = 0.001 AU at 280 nm

Analysis of p24 levels in SEC fractions of one RDpro-GFP sample by ELISA was used to determine which fractions contained the maximum vector particles. The detection range of the p24 ELISA used here is between 78 pg and 5 ng of p24 allowing the direct application of SEC fractions without an extra dilution or concentration step. The ELISA set up used here only

compares levels of p24 among fractions and cannot measure absolute levels of the protein in each of them. This was potentially due to two limitations in the early assay development. 1) The standard curve buffer contained 10% sheep serum and is normally used for crude samples that still contain FCS from DMEM used in vector production, whereas the SEC fractions are in TEN buffer without serum proteins, hence absolute OD_{405} readings of the samples were not comparable to absolute OD_{405} readings of the standard curve. 2) High concentration of Extravidin, an improved form of Streptavidin, resulted in a very fast enzymatic reaction after addition of the substrate p-nitrophenyl phosphate. The fast development did not allow reading high concentrated samples before the OD_{405} was out of detection range (>2.0 OD_{405}). Since this assay was only intended to identify the fractions that contained highest amount of p24, OD_{405} readings of SEC fractions relatively to each other were compared and no further experiments were conducted. P24 ELISA of RDpro-GFP shows that the highest amount of p24 protein is in the void peak (maximum in fraction 13, Figure 8).

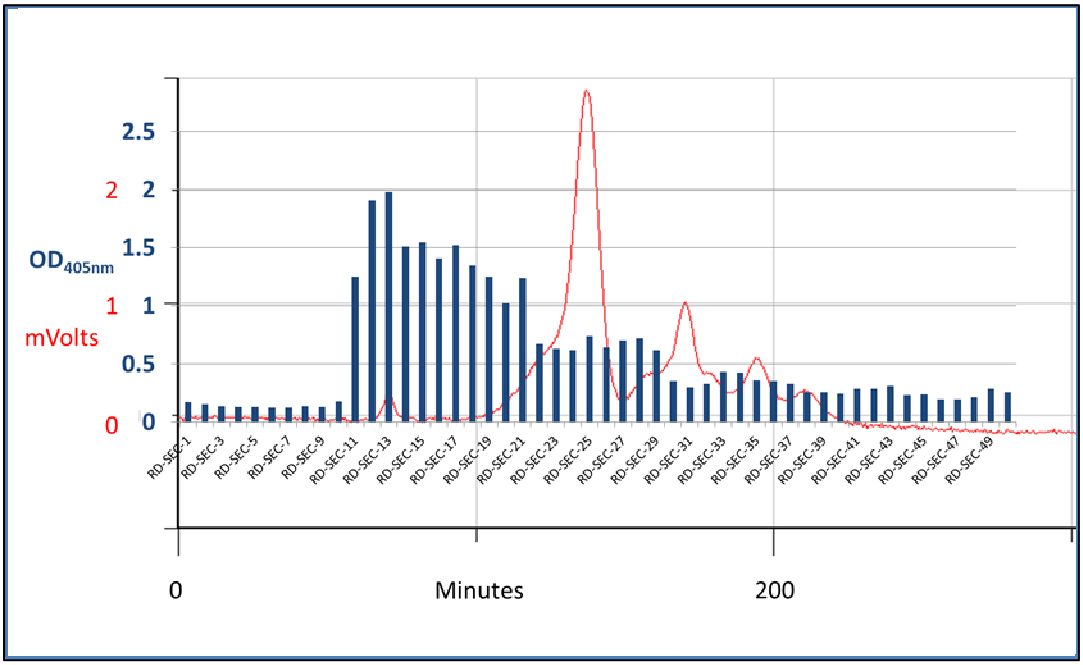


Figure 10: Maximum of p24 protein elutes in void peak fraction. Overlay of the SEC chromatogram of the analysed RDpro-GFP vector sample (900 µl, 40 fold concentrated, 1 mVolts = 0.001 AU at 280 nm, y-axis in red) and p24 ELISA OD_{405nm} readings (y-axis in blue) of each of the RDpro-GFP sample SEC fractions (RD-SEC, x-axis). Fraction 13, part of the void peak, shows highest OD_{405} reading of all analysed fractions. 0.8 ml/minutes in TEN buffer.

Elution of vector particles in the void peak was confirmed by Western blotting of the viral proteins. SEC fractions containing peak A to F were collected (Figure 9 A, peak A = void peak). Western blotting of viral proteins of these fractions showed detectable amounts of p24, VSV-G or RDpro only in the void peak of VSV-G-GFP and RD-GFP vector samples (peak A Figure 9B). In later eluting peaks (peaks B1 to F) no viral proteins were detected.

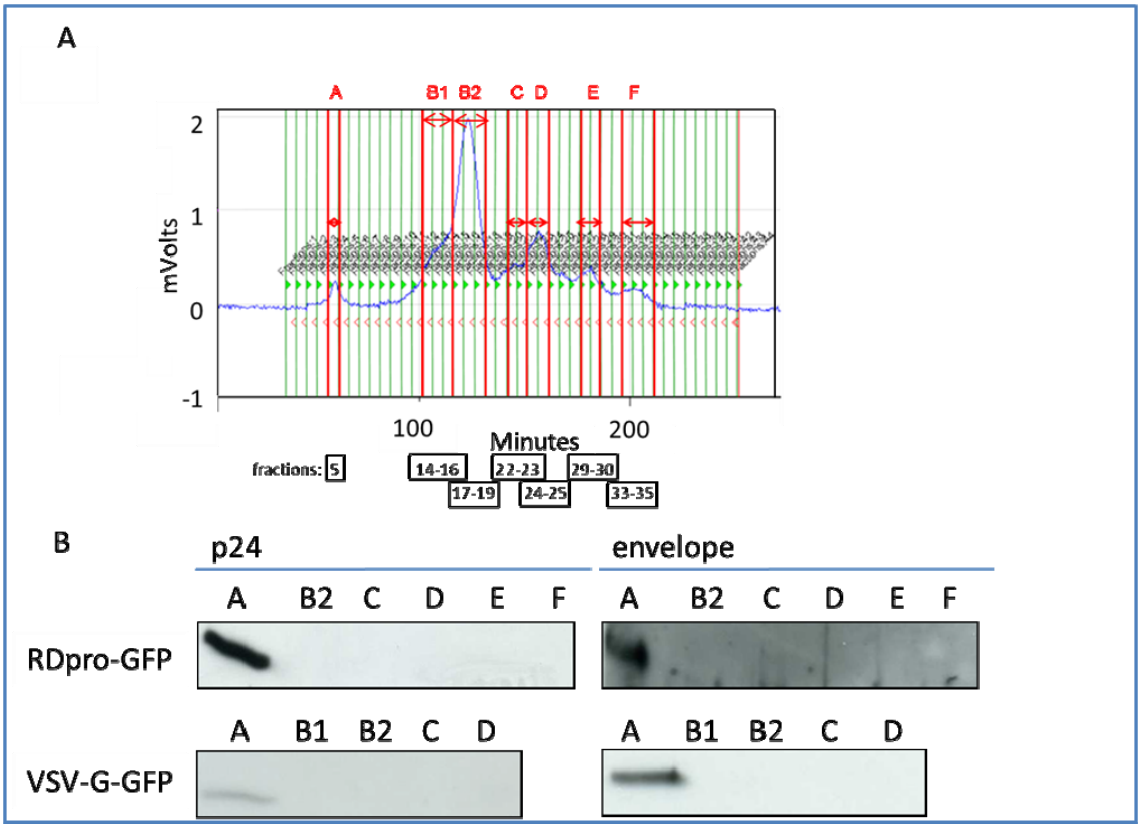


Figure 11: Viral proteins elute in void peak fraction. A) Representative SEC chromatogram of vector sample RDpro-GFP with labelled peaks A-F and fractions, B) Western blotting of viral proteins p24 and envelope proteins RDpro or VSV-G in different peak fractions; 1 mVolts = 0.001 AU at 280 nm. Signals for viral proteins were only detected in the void peak (peak A) fraction and not in later eluting protein peaks.

3.3.3. SDS-PAGE Silver Staining of Purified Vector Samples

For evaluation of LV purification by SEC, crude and purified vector samples were directly compared by silver staining of proteins after separation by SDS-PAGE. 2 µl of 900 µl crude samples were loaded per lane, equivalent to 0.22 % of total crude vector sample purified by one SEC run. In comparison 1 % of the 4 ml void peak fraction (40 µl) was applied. The silver stain in Figure 10 shows that dominant protein bands in SEC void peak fractions correlate with viral proteins sizes (57 kDa= VSV-G Env, 55 kDa = Gag precursor pr55Gag, 24 kDa = capsid/p24) and overall SEC fractions contain less protein. Many protein bands that can be seen in crude samples are not detectable in void peak fractions. With a detection limit of 0.8 ng of protein this method shows that these proteins have been significantly reduced.

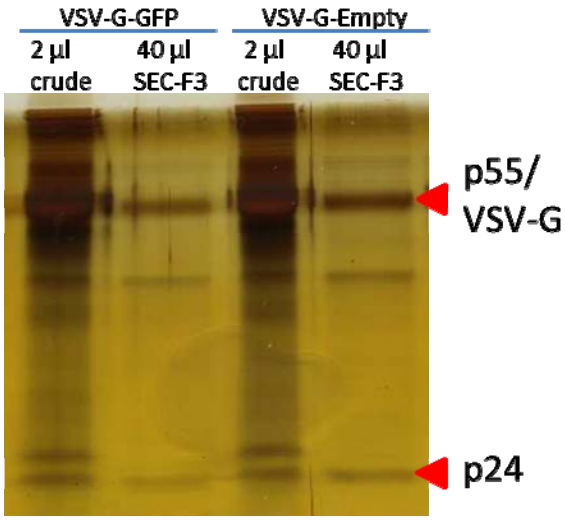


Figure 12: Vector purification by SEC. Silver Staining of SDS-PAGE of crude compared to SEC purified vector sample (0.22% of 900 µl crude sample loaded on SEC column, 1% of 4 ml void peak fraction). SEC fractions contain less protein, dominant bands are at the position of viral proteins pr55Gag (55 kDa), capsid p24 (24 kDa) and VSV-G Env (57 kDa).

3.3.4. Recovery of Infectious Particles in Void Peak

The transduction titer of void peak fractions was measured to assess if purified vector samples contain infectious vector particles. 293T cells were transduced with 0.45 µm filtered SEC void peak fractions of one VSV-G-GFP and one RDpro-GFP vector sample. The number of infected cells (GFP-positive cells) was counted by FACS. A total of 12% infectious vector particles could be recovered from crude samples after SEC purification in the analysed void peak (fractions 3 and 4) of the VSV-G-GFP vector sample (Figure 11 A & B, batch #78). Fractions 3 and 4 in the vector sample RDpro-GFP contained only 3% of the number of infectious particles in the crude sample loaded on the column (Figure 11 C, batch #95).

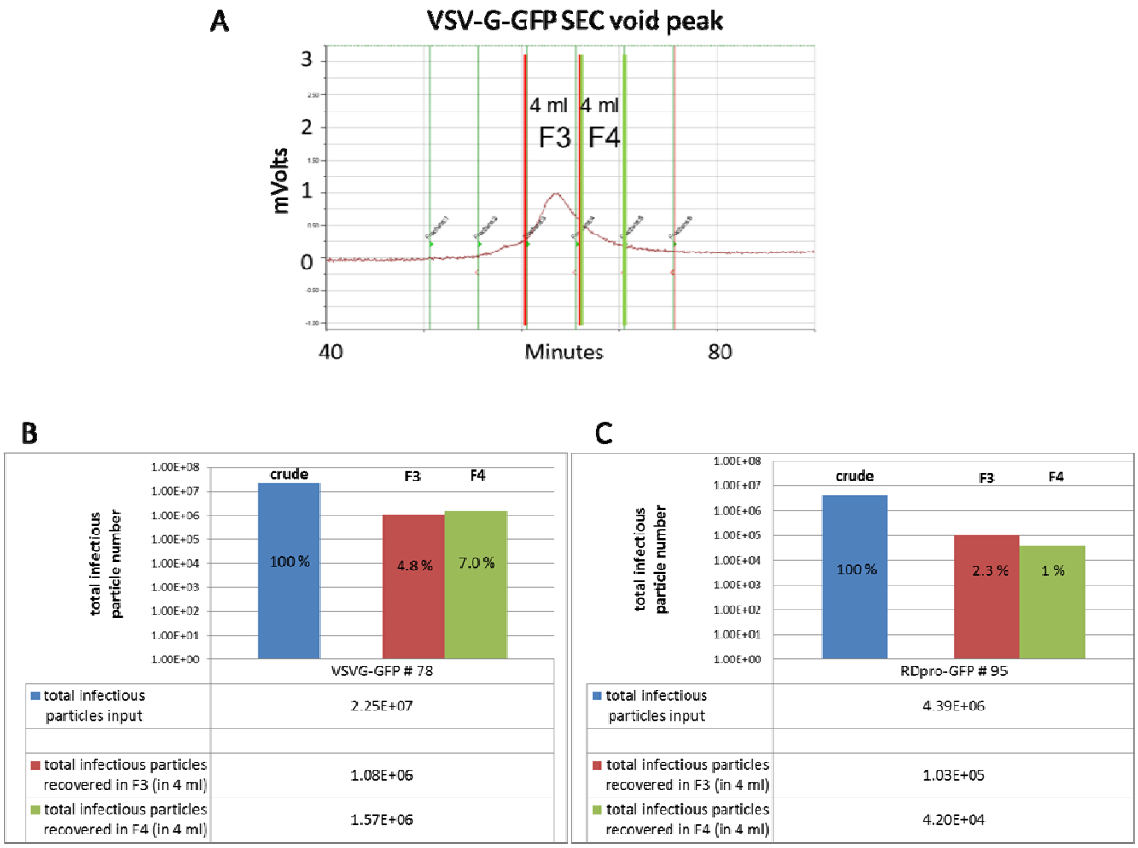


Figure 13: Recovery of infectious vector particles after SEC purification. A) SEC chromatogram of void peak of VSV-G-GFP vector sample, indicating analysed fractions for transfection titer, 1 mVolts = 0.001 AU at 280 nm, B) Percentage of total infectious vector particles recovered from SEC fractions 3 and 4 of samples VSV-G-GFP batch #78 and C) RDpro-GFP batch #95.

3.3.5. 240 fold Concentration of Stably Produced Samples Results in Comparable Amount of Viral Proteins to 40 fold Concentrated Transiently Produced Samples

Sample preparation of vector preparations for LC-MS/MS analysis involved the concentration of harvested cell supernatant by ultracentrifugation. Transiently and stably produced vector harvests were initially concentrated 40 fold. Compared to the void peak height of stably produced samples, void peaks of transiently produced samples were up to 10 fold higher when pseudotyped with VSV-G Env and up to 5 fold higher without an envelope protein (Figure 12 A). Semi-quantitative analysis by densitometry (see section 2.4.5) of p24 in void peak fractions of concentrated vectors showed that the viral protein is about six fold lower in stably produced vector samples compared to transiently produced samples (Figure 12 B). LC-MS/MS of 40 fold concentrated stably produced samples resulted in the detection of very few protein species (chapter 3, section 4.3.1.1). Even though stably produced samples were confirmed to contain viral proteins, the results of this first LC-MS/MS batch raised the concern that produced RDpro-pseudotyped samples did not contain sufficient numbers of viral particles to detect cellular proteins in low abundance and might not be comparable to transiently produced samples.

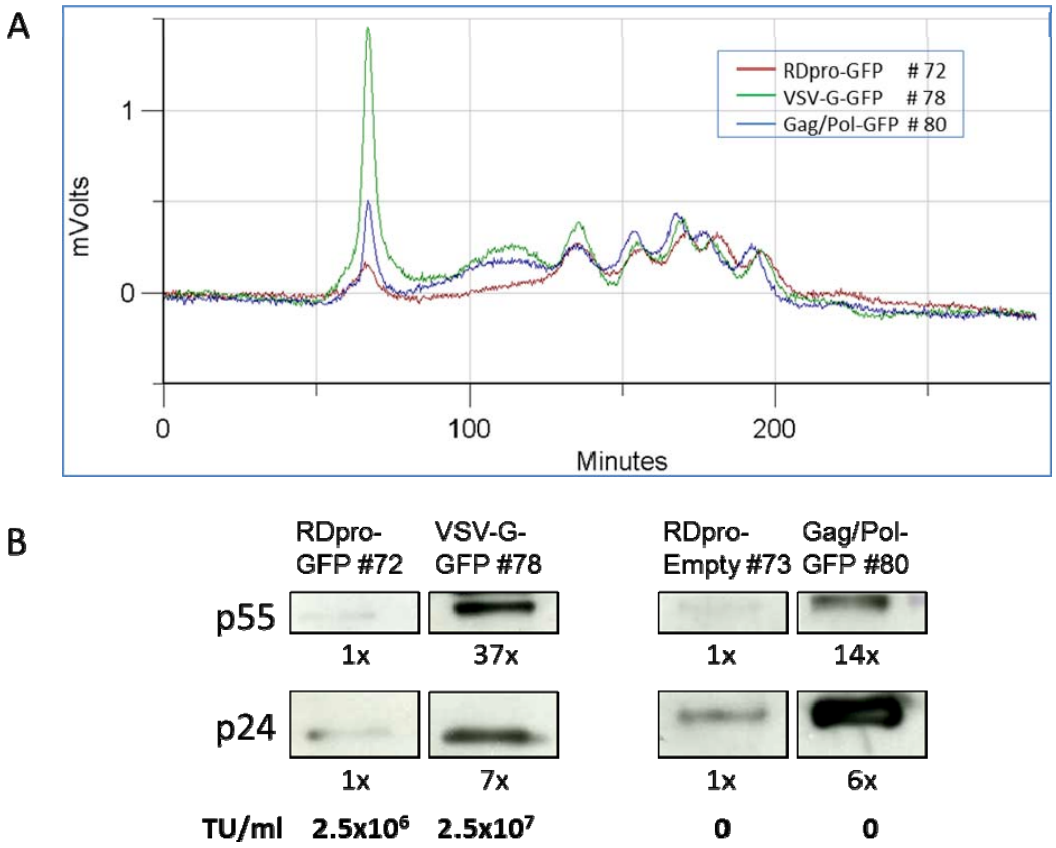


Figure 14: 40 fold concentrated stably produced vector samples contain 6 to7 fold lower amount of p24 than 40 fold concentrated transiently produced samples. A) Representative void peak height of SEC of 900 µl transiently and stably produced samples; SEC chromatograms of transiently produced samples result in 6 to 10 fold higher void peaks than of stably produced vectors, B) Western blotting of Gag (anti-p24) show semi-quantitative analysis by densitometry of viral protein signal strengths; levels of p24 in transiently produced samples are 6 to 7 fold higher (6x or 7x) and levels of GagPr55 14 to 37 fold higher compared to stably produced samples (1x) ; for protein detection void peak fractions were desalted and concentrated.

To gain comparable amounts of core protein p24 in transiently and stably produced samples vectors, hence comparable numbers of viral particles, vectors produced by STAR-RDpro-pHV and STAR-RDpro cells were concentrated 240 fold instead of 40 fold by ultracentrifugation (from now on referred to as 240 fold and 40 fold concentration). Purification of vectors using SEC showed an average of a 3 fold higher void peak (0.33-0.8 mVolts) compared to 40 fold concentrated stably produced samples (0.1-0.25 mVolts) (Figure 13 A). Western blotting of equivalent amounts of total protein per sample showed that 240 fold concentrated stably and 40 fold concentrated transiently produced samples contain comparable amounts of viral protein p24 (Figure 13 B). On average the transduction titer is 3 to 4 fold higher in 240 fold compared to 40 fold concentrated stably produced samples, hence infectivity of 240 fold concentrated RDpro-GFP vectors was still at least 3 fold lower compared to 40 fold VSV-G-GFP vectors (Figure 13 A). Void peak heights of VSV-G only 240 fold concentrated samples are about 6 fold higher than void peaks of 40 fold concentrated VSV-G pseudotyped vector samples (Figure 13 C). Purified vector samples were prepared for LC-MS/MS analysis.

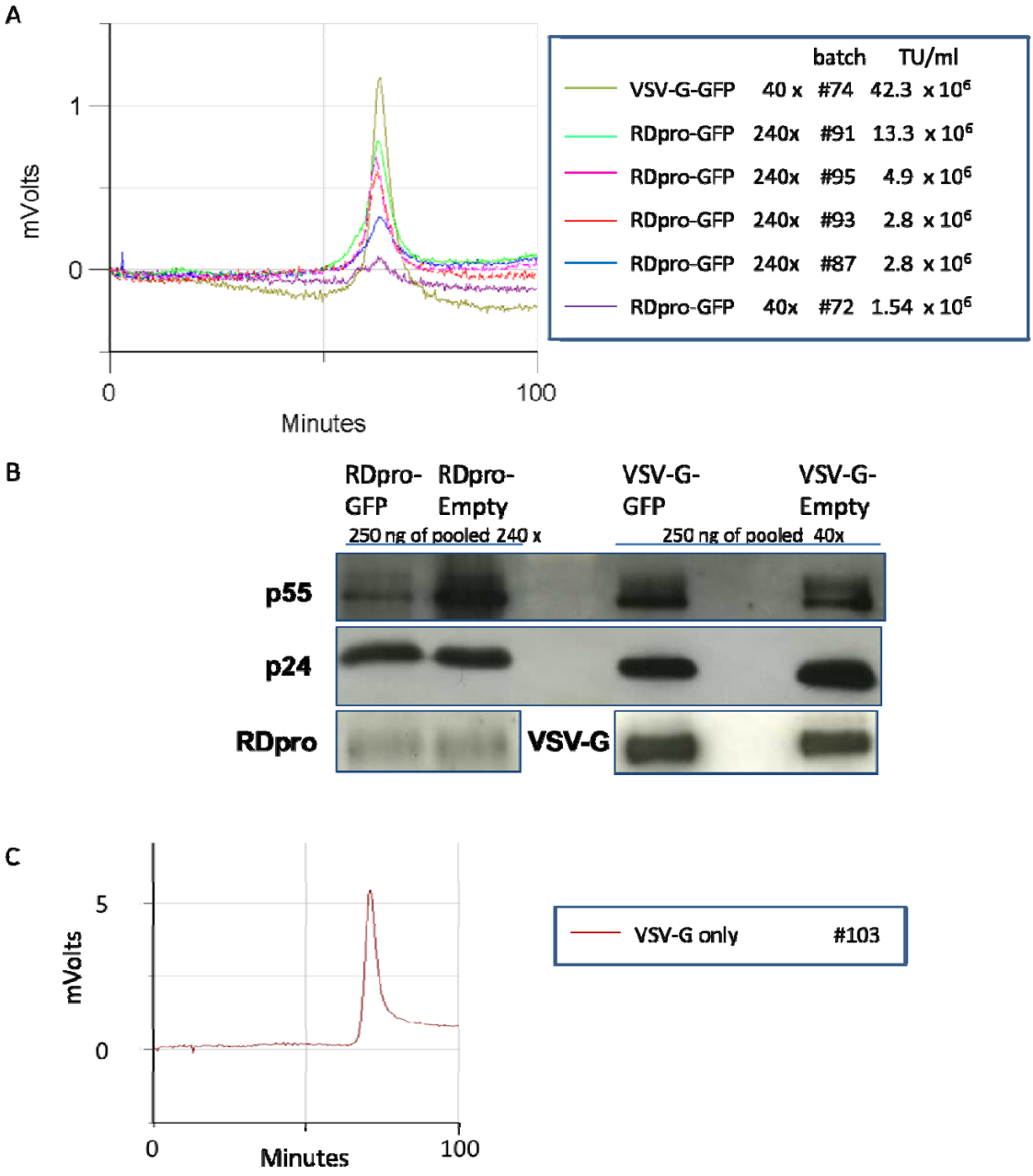


Figure 15: 240 fold concentrated stably produced vector samples contain similar viral protein amounts compared to 40 fold transiently produced vector samples. A) Representative void peak height of SEC of 900 µl 40x transiently and 40x/240x stably produced vector samples and transduction titer (TU/ml); void peak height average of 0.33-0.8 mVolts in 240 x compared to ~0.25 mVolts in 40 x concentrated stably produced sample; TU/ml increased 3 to 4 fold. B) Western blotting of Gag (anti-p24) and envelope protein RDpro and VSV-G show comparable amount of viral proteins when 250 ng total protein of 240 fold stably produced and 40 fold transiently produced samples are loaded. C) Void peak height of VSV-G only batch #103; 240 fold concentrated.

For comparison of viral and total protein content of purified transiently and stably produced vector samples for LC-MS/MS analysis, equivalent amounts of total protein of all vector samples were loaded on one acrylamide gel and visualised by Western blotting and silver staining. Western blotting of viral protein p24 shows that the processed Gag capsid protein p24 is the dominant form over the precursor protein pr55Gag in stably and transiently produced samples (Figure 13 A). Pr55Gag is not detectable in RDpro-Empty and transiently produced samples. A longer exposure time of the blot might reveal weak signals around the size of 55 kDa as was seen in previous Western blots of these samples (Figure 13 B). The control 'VSV-G only' does not contain p24 as expected which was confirmed later by LC-MS/MS (chapter 4). Presented in Figure 13 B is the Western blot of viral envelope proteins RD-pro and VSV-G. As expected the Gag/Pol-GFP control sample does not contain VSV-G protein.

Silver staining of these purified vector samples revealed only a few different bands in these samples which corresponded to the size of viral proteins identified by Western blot and confirm that dominant proteins in these samples are vector derived (Figure 14 C). Bands at the positions corresponding to the size of viral Env proteins RDpro (76 kDa) or VSV-G (57 kDa) respectively were detected only in samples that were expected to contain an Env protein, including RDpro-GFP as well as VSV-G-GFP, VSV-G-Empty and VSV-G only. In accordance with these results no signal at the position of 76 or 57 kDa was detected at that position in Gag/Pol-GFP vector sample confirming it does not contain an Env protein. RDpro Env in RDpro-Empty samples was below detection limit using silver staining, which is 0.8 ng total protein per lane. Silver staining of stably produced samples shows a weak signal at position equivalent to 76 kDa (Figure 13 C), suggesting RDpro envelope protein is of lower concentration compared to VSV-G-envelope protein in transiently produced samples. Signals at position of 57 kDa are the most dominant band in VSV-G pseudotyped samples and are likely to be the VSV-G envelope protein. However pr55Gag might contribute to that signal as it is also visible in silver stained stably produced samples (Figure 13 C and Figure 15). The capsid protein p24 was detected at the position of 24kDa in stably and transiently produced samples and was absent in VSV-G only

samples as expected. Silver staining of lower amounts of samples (Figure 15; 50 ng instead of 100ng total protein/lane) confirm that dominant proteins correspond to the size of viral proteins with p24 being the most abundant protein in stably produced purified vector samples.

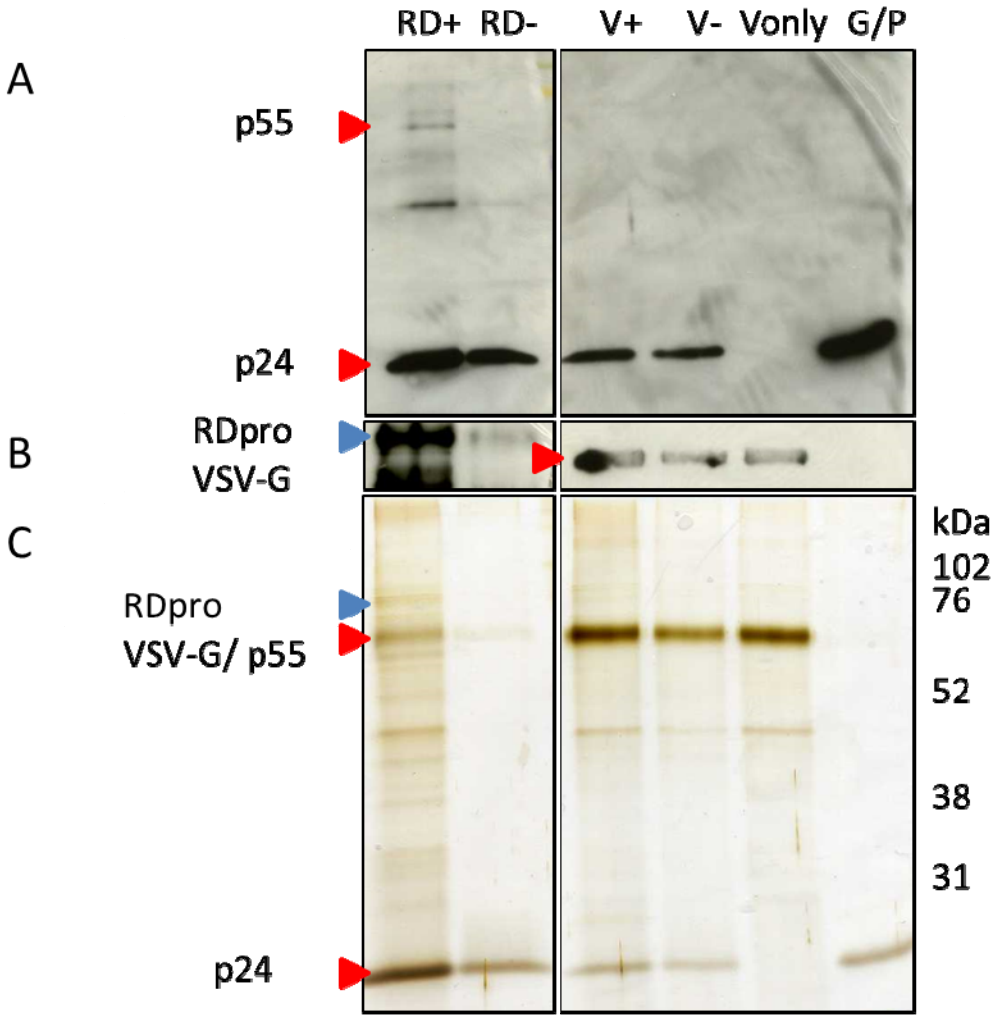


Figure 16: Purification by SEC of all five vector samples and VSV-G control. A) Western Blotting of SEC-purified vector sample (250ng/lane) shows p24 in all LV samples and not in 'VSV-G only' control as expected. Less pr55Gag is detected in RDpro-Empty samples compared to RDpro-GFP and transiently produced samples, B) RDpro and VSV-G protein Western blotting of SEC-purified vector sample (250ng/lane). A) and B) Western blotting confirms that vector samples eluted in SEC void peak containing mainly viral proteins, shown by silver stain. C) Silver Stain of

SEC-purified vector sample (100ng/lane), dominant bands correlate with viral protein sizes. RD+=RDpro-GFP, RD-=RDpro-Empty, V+=VSV-G-GFP, V-=VSV-G-Empty, Vonly=VSV-G-only, G/P=Gag-Pol-GFP.

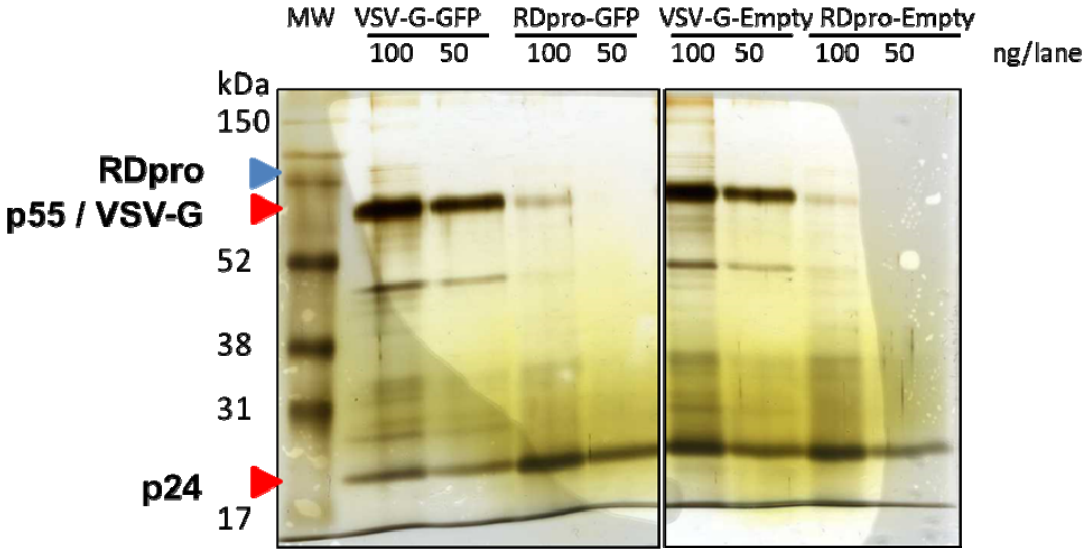


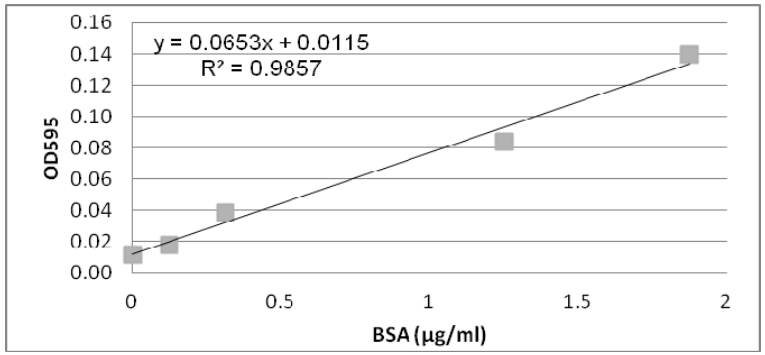
Figure 17: Comparable amounts of p24 in vector samples for LC-MS/MS. Silver staining of SDS-Page of vector samples after SEC purification and preparation for LC-MS/MS analysis. Transiently and stably produced samples show dominant bands at the position of pr55Gag (55 kDa) and capsid p24 (24 kDa), the latter being of comparable signal strength. Signals at position equivalent of 57 kDa is stronger in transiently produced samples, correlating to the size of VSV-G Env, whereas RDpro Env (76 kDa) is below detection limit. (V+ = VSV-G-GFP, V- = VSV-G-Empty, RD+ = RDpro-GFP, RD- = RDpro-Empty); 100 or 50 ng total protein/lane loaded as indicated.

3.3.6. Desalting and Lyophilisation of SEC Void Fractions in Preparation for MS study

Purified vector samples in SEC void peak fractions could not be analysed directly by LC-MS/MS. This is because proteins in vector samples needed to be digested by trypsin under the condition of pH 8.0. SEC fractions were eluted in TEN buffer (150 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl) that had a pH of 7.4, thus was too low for adequate protein digestion by trypsin. For each vector sample between eight and sixteen void peak fractions were pooled and dialysed to exchange TEN buffer (pH 7.4) with 10 mM ammonium bicarbonate (ABC, pH 8.5). ABC is volatile and sublimes into ammonium, water and carbon dioxide making it attractive for buffer exchange when the protein sample needs to be salt free and without any additional components added by dialysis which would stay behind after lyophilisation (Overcashier et al. 1997). Assuming complete equilibration after dialysis over night the salt concentration of sample was reduced 375 times in ABC at that stage resulting in a salt molarity of 0.4 mM NaCl, 0.0026 mM EDTA and 0.026 mM Tris-HCL. The buffer was changed twice the following day and samples dialysed for at least 2 hours after each buffer change. This resulted in a total salt dilution factor of $52.7x10^6$. A minimum of 10 µg of total protein in a small volume of up to 500 µl was needed for each vector samples for a LC-MS/MS run. To reduce sample volume dialysed fractions were freeze dried and resuspended in water. To quantify the total proteins in samples for LC-MS/MS analysis, an assay needed to be developed that allowed measuring low concentrations of total protein in a small sample volume. The standard Bradford assay was modified by increasing the sample to dye ratio. An aliquot of 5 µl of sample was diluted 1:160 in water before addition of 200 µl concentrated Bradford dye. Furthermore the sample was acidified by adding 10 µl of 0.1 N HCl helping protein binding to Coomassie® Brilliant Blue G-250 (Ramagli and Rodriguez, 1985). Assay details for the measurement of vector samples MS set 2 are shown in Table 11. After protein quantification samples were processed in the proteomics facility for LC-MS/MS analysis.

Table 11: Bradford Assay of Purified Vector Sample Set 2 for LC-MS/MS.

Standard curve					OD595 reading 1	OD595 reading 2	OD595 reading 3	OD595 AVERAGE	Protein (µg/ml)	Total protein (µg)
BSA concentration (µg/ml)	volume of BSA (50 µg/ml, µl)	10 mM ABC (µl)	0.1N HCl (µl)	H_2O (µl)					diluted 1:160 (5 µl in 800 µl)	in 200 µl
0	0	5	10	785	0.0106	0.0119		0.0113		
0.125	2	5	10	783	0.0117	0.0238		0.0178		
0.3125	5	5	10	780	0.0346	0.0419		0.0383		
1.25	20	5	10	765	0.0836	0.0909		0.0836		
1.875	30	5	10	755	0.1378	0.1410		0.1394		
samples		sample (µl)								
RDpro-GFP		5	10	785	0.0171	0.0219	0.0172	0.0187	0.11	**3.52**
RDpro-Empty		5	10	785	0.0107	0.0149	0.0197	0.0151	0.06	**1.92**
VSV-G-GFP		5	10	785	0.0369	0.0454	0.0324	0.0382	0.41	**13.12**
VSV-G-Empty		5	10	785	0.0252	0.0338	0.0347	0.0312	0.30	**9.60**



Modified Bradford assay can detect low amounts of protein (0.125-1.875 µg/ml total protein). Total protein in 200 µl of dialysed and freeze dried SEC purified from 7.2 ml crude vector samples. The here developed modified Bradford assay allowed to use only 5 µl of prepared vector sample.

3.3.7. Summary of Vector Sample Preparation for LC-MS/MS

In total three sets of transiently and stably produced vector samples were prepared for LC-MS/MS analysis. A summary of the sample preparation process is outlined in Figure 16. Initially MS sets 1 and 2 of the five vector samples (VSV-G-only was only analysed in MS set 3) were produced by 40 fold concentration using ultracentrifugation. The amount of total protein was below 5 µg for RDpro-GFP of set 1 and 2 as well as for RDpro-Empty set 2 samples. As described in the following chapter (chapter 4) only very few different protein species were detected in stably produced samples. In order to analyse equivalent amounts of vector particles by LC-MS/MS, the amounts of viral protein p24 in stably and transiently produced samples was normalised by concentrating stably produced samples 240 fold, resulting in MS set 3 of stably produced samples (Table 12). Only the results from a comparable number of viral particles in stably and transiently produced samples allowed a meaningful comparison of the number of different cellular protein species in samples produced by different production methods and pseudotyped with different envelope proteins. Furthermore including another vector sample, transiently produced RDpro pseudotyped vectors, would could have allowed distinguishing between vector core or viral envelope associated proteins.

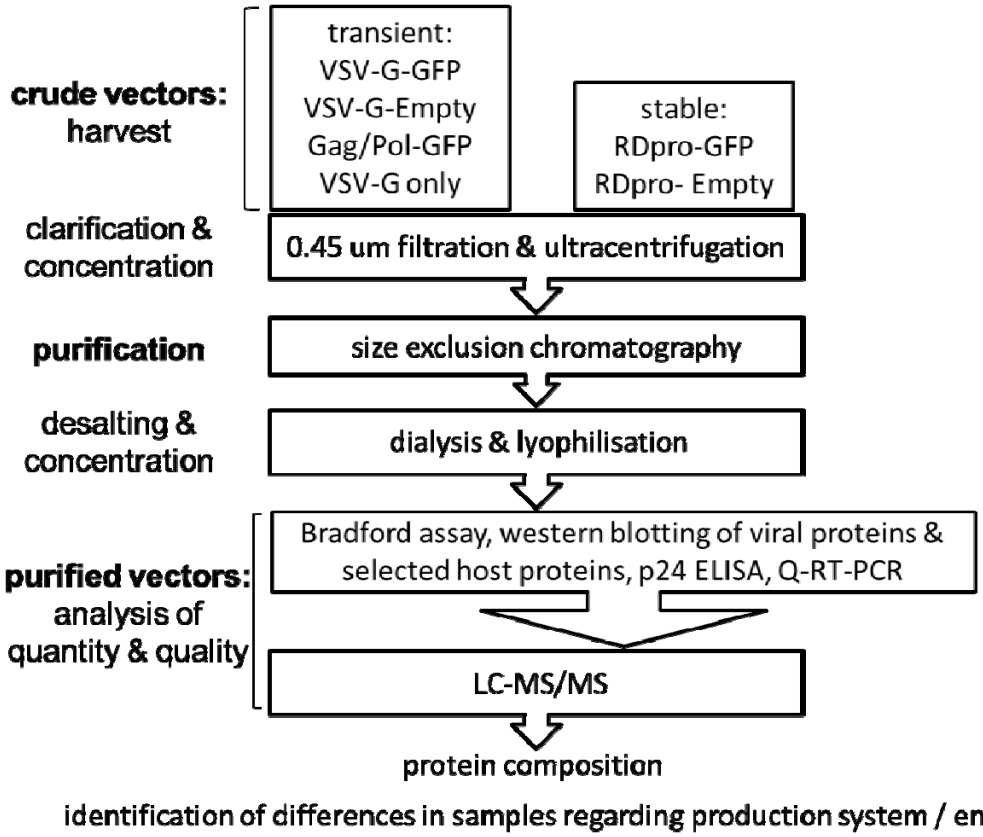


Figure 18: Vector Sample Processing for LC-MS/MS Analysis: Vectors were harvested in cell culture medium and processed in several steps before viral and host protein detection by LC-MS/MS.

Table 12: Overview of Vector Samples Prepared for Mass Spectrometry (MS) analysis

vector sample	MS set	40 x crude sample (ml)	240 x crude sample (ml)	total input of transducing units ($x10^6$ TU/ml)	pooled void fractions (4ml x)	total protein (µg)	total protein used for MS (µg)
VSV-G-GFP	set 1	7.2	-	238.7	8	8.6	**8.6**
VSV-G-Empty		7.2	-	-	8	16.6	**8.6**
Gag/Pol-GFP		7.2	-	-	8	11.7	**8.6**
VSV-G-GFP	set 2	7.2	-	302.4	8	12.4	**12.4**
VSV-G-Empty		7.2	-	-	8	9.1	**9.1**
VSV-G-GFP	set 3	13.5	-	1314.9	15	76.5	**10**
VSV-G-Empty		13.5	-	-	15	67.1	**10**
Gag/Pol-GFP		12.6	-	ND	16	34.7	**10**
VSV-G-only	set 1	-	7.0	-	8	138.2	**10.0**
RDpro-GFP	set 1	7.2	-	6.5	8	2.5	**2.5**
RDpro-Empty		7.2	-	-	8	10.4	**10.4**
RDpro-GFP	set 2	7.2	-	15.1	8	3.4	**3.4**
RDpro-Empty		7.2	-	-	8	1.7	**1.7**
RDpro-GFP	set 3	-	10.8	50.9	12	20.8	**10.0**
RDpro-Empty		-	10.8	-	12	58.2	**10.0**

Three sets of each transiently and stably produced samples were analysed by LC-MS/MS. MS sets for stably produced samples include 40 (MS set 1 and 2) and 240 fold (MS set 3) concentrated preparations; whereas transiently produced vector samples included only 40 fold concentrated vector samples (MS set 1, 2 and 3; VSV-G control 240 fold concentrated) The total amount of protein used for LC-MS/MS analysis is shown. One void peak fraction (4ml) was derived from 900 µl crude vector harvest, which is derived from 18 ml of crude unconcentrated vector harvest from one 15 cm dish (see section 2.3.1).

3.4. Discussion

Growing numbers of clinical trials for larger patient cohorts require large scale vector production which is difficult to achieve with the transient transfection methods currently used in clinical protocols. The packaging cell line STAR (Ikeda et al., 2003) uses RDpro as an alternative envelope to VSV-G. In comparison to VSV-G, RDpro is not as toxic to producer cells, it is more resistant to inactivation by human serum (Strang et al., 2004) and its receptor is abundant on HSCs. The aim of this part of the study was to assess production and purification of LVs using the two different vector production methods for subsequent analysis by mass spectrometry (LC-MS/MS) followed by comparison of identified vector associated cellular proteins (chapter 4). Transiently produced samples included two VSV-G-pseudotyped samples (with or without a vector genome; VSV-G-GFP and VSV-G-Empty) as well as particles without RDpro or VSV-G envelope protein (Gag/Pol-GFP). Furthermore a sample without HIV-1 proteins was prepared (VSV-G-only). Stably produced samples were RDpro-pseudotyped vectors with and without a vector genome produced by the packaging cell line STAR (Ikeda et al., 2003).

Analysis of vector samples showed that the transduction titer of crude samples was 10 fold higher in 40 fold concentrated transiently compared to 40 fold concentrated stably prepared samples. Western blotting showed that purified vector samples produced by STAR cells contained about 6 fold lower amounts of viral protein p24 compared to VSV-G pseudotyped transiently produced samples packaged by a comparable number of producer cells. Increasing the concentration factor of crude samples to 240 fold resulted in comparable p24 levels in vector samples. However these vectors were still at least 3 fold less infectious than VSV-G-pseudotyped particles. Higher infectivity of VSV-G LVs in comparison to RDpro-pseudotyped LVs has also been reported in other studies (Ikeda et al., 2003, Strang et al., 2004, Bell et al., 2010).

After concentration of stably produced samples was increased 6 fold to a total concentration factor of 240, p24 levels in purified vectors increased 6 fold, comparable to p24 levels in

transiently produced samples. However the transduction titer of 240 fold concentrated RD was only increased by 3 fold, indicating that the 240 fold concentration resulted in higher level of non-infectious vectors. This may potentially be due to shedding of RDpro Env during concentration, thus a high concentration factor is possibly not suitable for vector particles pseudotyped with RDpro Env. Up-scaling of vector production by increasing the concentration factor is relatively easy due to the progress made in vector production process development of stable producer cell lines, such as their adaption to grow in suspension culture in large bioreactors (Ghani et al., 2006, Broussau et al., 2008). Concentration factors of 200 (Merten et al., 2011) and up to 2000 (Schweizer and Merten, 2010) have been used in vector production. High concentration factors might affect some envelope proteins more than others. It needs to be ensured that vectors are not damaged during preparation for clinical application resulting in low or non-infectious particles to avoid administration of high doses of viral proteins.

Various methods have been employed to perform vector purification during vector processing for their clinical application. In up-scaled manufacturing protocols of clinical trials anion-exchange chromatography (AExc) has been used to purify HIV-1 derived (Schweizer and Merten, 2010, Merten et al., 2011) and EIAV derived LVs (Stewart et al., 2010). Publications of recent clinical trials describe similar vector production processes (Aiuti et al., 2013, Biffi et al., 2013). Vectors are clarified by filtration and purified by diethylaminoethyl (DEAE) AExc which is based on the purification of negatively charged proteins including vector particles by a positively charged chromatographic support and subsequent elution with high salt buffer at increasing ionic strength. Also referred to as gel filtration, SEC can follow and is used for final removal of trace contaminants such as smaller proteins with a similar negative charge than vector particles and for reformulation in storage buffer including adjustments of pH (Aiuti et al., 2013, Biffi et al., 2013). In comparison to AExc SEC has been shown to be less scalable due to its dilution effect; therefore it is less suitable as the first step in downstream vector processing however it can be used for vector purification in small scale experiments with low samples volumes (Slepushkin et al., 2003, Levine et al. 2006).

Purification of LVs by SEC has been used in our laboratory (Keating et al., 2008). Others have used process sequences of clarification and concentration by ultrafiltration of harvested vectors followed by SEC purification without an additional purification step (Slepushkin et al., 2003, Levine et al., 2006). Sephacryl S-500 medium has been applied for purification of VSV-G pseudotyped LV products, VRX496, used in a small scale clinical trial for the treatment of HIV positive individuals. Transfiguracion et al. (2003) have assessed purification results of SEC by analytical AExc showing that the main non-viral protein contaminants in crude samples were BSA and low molecular proteins were not detectable in purified samples. Our results showed that void peak fractions contained vector particles as demonstrated by ELISA, Western blotting and detection of infectious particles. Levels of p24 protein were the highest in void peak fractions relative to all other fractions collected over the entire length of the chromatography run. By Western blotting viral proteins p24 and envelope RDpro as well as VSV-G were only detected in fractions corresponding to the void peak and not in fractions of later eluting peaks. This suggests that vector particles are of comparable mass to wild type HIV-1 virions of 277 MDa (Carlson et al., 2008) by eluting at the expected time in the void peak. Furthermore results show that Sephacryl 500-HR SEC media is suitable for separation of vectors from molecular weight contaminants with a size of $<2x10^7$ M_r.

Identification of potential true differences in vector associated host proteins among vector samples can only be ensured when possible co-purification of comparable sized proteins or protein aggregates with vectors is prevented as they could be identified as false-positive vector associated proteins. Tubulovesicular structures, similar to microvesicles, have been found in concentrated LVs pseudotyped with VSV-G (Pichlmair et al., 2007). The vector sample, Gag/Pol-GFP, did not carry VSV-G Env and was prepared to distinguish cellular proteins associated with vectors from VSV-G vesicle bound proteins in MS samples. Gag/Pol-GFP vector samples were produced by transiently transfection of the Gag-Pol packaging plasmid and GFP transfer vector (pRRLSIN.cPPT.PGK-GFP.WPRE) into 293T cells without any envelope encoding plasmid. Pr55Gag can assemble into virus like particles (VLPs) in the absence of envelope

proteins in cells and secrete extracellular (Gheysen, 1989). In our experiments the Gag/Pol encoding plasmid was transfected into cells, hence HIV-1 protease was expected to be expressed cleaving Pr55Gag allowing the detection of capsid p24 protein in vector samples. Detection of similar levels of p24 in purified Gag/Pol-GFP samples as well as a similar void peak height compared to other transiently produced samples indicated that LV sample Gag/Pol-GFP contained comparable numbers of viral particles without an envelope protein. Furthermore a sample, 'VSV-G only', was prepared by transfection of the VSV-G Env plasmid and GFP transfer vector plasmid and did not express HIV-1 structural proteins in order to investigate which proteins could be associated with VSV-G Env itself. Void peak fractions of this sample were anticipated to contain less protein than samples containing vector particles; hence the VSV-G only sample was concentrated 240 fold. Western blotting of the VSV-G-only sample showed similar levels of VSV-G Env to VSV-G-GFP and VSV-G-Empty samples, indicating transiently produced VSV-G pseudotyped samples contained similar levels of VSV-G complexes, such as aggregates and vesicles.

Future investigations can be undertaken regarding the removal of impurities closely related to the size and shape of functional LVs, such as non-viral membrane vesicles. Another purification step could be included that can separate viral particles from vesicles. Rate zonal ultracentrifugation using iodixanol, a density gradient medium, has been used to eliminate cell membrane vesicles during purification of MoMLV-derived retroviral particles (Segura et al., 2006a). In order to assess if any contaminating vesicles are present electron microscopy of crude and purified vector particles can be performed (Transfiguracion et al., 2003, Pichlmair et al., 2007). Further control samples can be included, such as harvests from mock-transfected 293T cells, transfected with DNA plasmids not carrying HIV-1 or VSV-G sequences. In addition vectors could be produced in serum free medium, using for example, Opti-MEM, already from 24 hours post-transfection (see section 2.3.1).

Vector purity was analysed using silver staining. During silver staining of polyacrylamide gels separated vector samples, silver ions are reduced to free metallic silver after forming complexes with a biochemical group of amino acids or nucleic acids (Merril and Pratt, 1986, Switzer C Robert, 1979). Silver stains showed that dominant bands of concentrated and purified vector samples correlated with the size of viral proteins detected by Western blotting. Purified samples contained fewer different protein species as many protein bands detected in crude samples could not be detected in purified vector samples.

Fully functional vector particles, up to 12% of crude VSV-G pseudotyped vectors, could be recovered in SEC void peak fractions. Recovery of infectious particles of RDpro-GFP samples was lower and may be due to low numbers of infectious vector particles in crude samples and lower stability of the RDpro Env protein compared to VSV-G Env. Detection of infectious particles in void peak fractions confirms results from p24 ELISA and Western blotting, showing elution of vector particles in the void peak with some of them being mature and infectious. A recovery of up to 80% of infectious particles of purified transiently produced VSV-G pseudotyped LVs by SEC with Sephacryl S-500 medium has been reported previously (Slepushkin et al., 2003). Transfiguracion et al. (2003) separated VSV-G pseudotyped transiently produced LVs from smaller contaminants with a Sepharose CL-4B gel matrix using the same buffer as or study, TEN buffer (chapter 2), but a lower flow rate of 0.4 ml/minutes. Several fractions were titered including void peak fractions as well as fractions before and after the void peak showing that all are containing infectious particles with the majority of infectious particles recovered in the void peak fraction. Infectious particles in all measured fractions added up to an overall recovery of 70% of infectious viral particles loaded on the SEC column. Here only two fractions were titered, the void peak and the next fraction eluting after the void peak. Later eluting fractions could contain infectious particles hence the percentage of overall recovered infectious vectors may be higher than 12% in transiently produced samples. However the aim was to collect vector containing fractions without compromising purity of vector particles. Quantification of overall recovery was not relevant for protein analysis of

vector particles, hence only the void peak fraction containing highest levels of particles was collected. On the other hand it needs to be taken into consideration that a reduced number of infectious particles could be due to vector particle damage or dissociation of vector-associated host cell proteins that mediate infectivity. This could also apply to host proteins that are vector associated due to a function in vector assembly and budding which could consequently be lost for subsequent LC-MS/MS analysis. An alternative SEC buffer could be used such as PBS or serum free cell culture medium used in gel filtration steps in vector preparation for clinical trials (Aiuti et al., 2013). Prior to SEC AExc is applied in these trials. AEx is either based on chromatography medium such as diethylaminoethyl (DEAE) or membranes (Mustang Q). AEx using Mustang Q membranes showed a slightly lower recovery of infectious titers (65%) in comparison to SEC (80%) (Slepushkin et al., 2003). Alternatively a gradually increasing salt concentration in the elution buffer allows dissociation and elution of vector particles. It was shown that vectors eluting at a high salt concentration have a lower transduction efficiency compared to vectors eluting at a lower salt concentration during the same chromatography run (Yamada et al., 2003). The high amount of salt in the elution buffer potentially damages vector particles by stripping off surface proteins. Generally though, for large scale purification purposes AEXc is more suitable than SEC as it does not require initial concentration of vector harvests due to high flow rates.

Dialysis and lyophilisation were used to desalt and further concentrate vectors since the identification of host and viral proteins in purified vector samples by LC-MS/MS analysis did not require functional particles. In an attempt to retain all associated proteins in the purified vector samples a dialysis cassette with a low molecular weight cut- off of 10 kDa was used. In large scale vector production concentration of purified vector samples is normally carried out before re-formulation by SEC (Aiuti et al. 2013; Biffi et al. 2013). Merten et al. (2011) applied hollow fibre tangenital flow requiring a relatively complex set up compared to dialysis and lyophilisation when using for small scale vector purification.

The here presented results outline LVs preparation by two different production systems followed by a suitable method of vector purification for comparative analysis of vector associated host-protein. LC-MS/MS analysis results are described in the following chapter.

4. Protein Composition of Purified Vector Samples Analysed by Tandem Mass Spectrometry

4.1. Introduction

Lentiviral vectors (LVs) analysed in this study are based on HIV-1 group M subtype B isolate HXB2. The vectors contain structural proteins encoded by *gag* and *pol* as well as a vector genome, and are pseudotyped with either VSV-G Env or RDpro. Like most viruses, the replication of HIV-1 and packaging of HIV-1 derived vectors relies heavily on host cell proteins that interact with viral proteins during its replication or production in packaging cells, respectively. Several cellular proteins have been identified to interact with viral proteins during wild-type HIV-1 assembly. Studying the role of cellular proteins associated with purified wild type HIV-1, such as clathrin, has elucidated the involvement of host proteins in virus assembly and budding (Ott, 2008, Chertova et al., 2006). Clathrin-associated heterotetrameric adaptor protein (AP) complexes are involved in cellular cargo sorting and transport (Ohno 2006). In HIV-1 budding AP-1 and AP-2 bind to the C-terminus of HIV-1 Env subunit gp41 regulating its location to the virus budding site (Berlioz-Torrent et al., 1999). RNA interference (RNAi) induced AP-1u silencing showed a reduced Gag release from HIV-1 infected cells (Camus et al., 2007).

Previous studies have shown that purified lentiviral vectors incorporate or bind to cellular proteins, such as actin, heat shock 70kDa protein 1A/1B (HSPA1A) and programmed cell death 6-interacting protein (ALIX) (Denard et al., 2009, Wheeler et al., 2007). Several groups (Wheeler et al., 2007, Segura et al., 2008, Denard et al., 2009) have used mass spectrometry (MS) to identify vector associated host cell proteins. Mass Spectrometry permits the comprehensive analysis of the protein composition of complex biological samples and is a well-established technique. Proteins of a sample can be analysed by mapping of the detected peptides and peptide sequencing using tandem MS (MS/MS). MS/MS involves the

measurement of peptide mass by a mass spectrometer followed by fragmentation of the peptides using electrospray and a second mass spectrometer to identify the amino acid sequence. Peptide preparation can be achieved by two different approaches. Single protein analysis is based on separation of proteins of a sample in a two dimensional SDS-PAGE (2D-PAGE) with subsequent excision of protein spots and digestion using a protease, commonly trypsin. In 2D-PAGE proteins are separated in one dimension by size and in the second dimension by isoelectric point (pI). This method was used to analyse the *Haemophilus influenzae* proteome identifying 502 different proteins (Langen et al., 2000). Also referred to as shotgun analysis, LC-MS/MS allows identification of a complex protein mixture. For analysis of complex samples two-dimensional LC-MS/MS (LC/LC-MS/MS) can be used, also called multidimensional protein identification technology (MudPIT) which is based on peptide separation by columns packed with a strong cation exchanger (SCX) and a reverse-phase (RP) liquid chromatography (Link et al., 1999). This method was used for the analysis of the *Saccharomyces cerevisiae* proteome detecting 1485 proteins (Washburn et al., 2001).

For analysis of size exclusion chromatography (SEC) purified lentiviral vectors we use LC-MS/MS, for which purified vector samples are trypsinised and resulting peptides desalted using reverse-phase chromatography columns (see chapter 2). LC-MS/MS has been used in previous studies of proteome analysis of purified virions such as HIV-1 (Chertova et al., 2006), Kaposi's Sarcoma-Associated Herpesvirus (KSHV) (Zhu et al., 2005) and Epstein-Barr virus (Johannsen et al., 2004).

4.2. Aim

The aim of this chapter is to analyse five different types of SEC-purified lentiviral vectors, as well as the VSV-G-only control sample, by LC-MS/MS to identify their associated cellular proteins and determine any similarities or differences. Vector samples include VSV-G pseudotyped transiently produced vector samples as well as RDpro-pseudotyped stably produced vectors. Vectors with and without a vector genome were analysed as well as a sample prepared by transient transfection of only VSV-G Env with the vector genome DNA plasmid used as carrier DNA. Differences and similarities in host cell protein composition among vector samples are described that are the basis for further analysis regarding their function in vector assembly and budding from the producer cell.

4.3. Results

4.3.1. Optimisation of Vector Sample Preparation for MS analysis

4.3.1.1. Sample Optimisation to increase Protein Quantity

Initially two vector samples were analysed by LC-MS/MS including stably produced RDpro-GFP and transiently produced VSV-G-GFP. Void peak fractions from 7 SEC runs of 40-fold concentrated RDpro-GFP were pooled. Void peak fractions from 5 SEC runs of VSV-G-GFP were pooled and processed for mass spectrometry analysis. In stably produced samples only Gag protein was identified. In VSV-G-GFP vectors an additional six human and seven bovine proteins as well as VSV-G Env were detected.

For MS analysis of the next set of vector samples (referred to as 'MS set 1' below) the total amount of proteins was increased to a level that could be detected by modified Bradford protein assay by pooling void peak fractions from 8 SEC runs, instead of 5 SEC runs, of each sample and processing as previously described (see chapter 3, Table 3). MS set 1 included five LV vector samples analysed by LC-MS/MS, namely VSV-G-GFP, VSV-G-Empty, Gag/Pol-GFP, RDpro-GFP and RDpro-Empty. RDpro-GFP and RDpro-Empty samples, produced by STAR cells, contained 2.5 µg and 10.4 µg total protein, respectively. Transiently produced samples contained between 8.6 and 16.6 µg total protein. A total protein amount of 8.6 µg of each transiently produced sample was analysed. The number of different human cellular protein species in transiently produced samples identified by MS ranged from 14 to 49 proteins whereas stably produced samples contained only 1 to 5 proteins.

Results of the subsequently analysed MS set 2 were similar, with stably produced samples containing significantly lower levels of total protein after pooling of SEC void peak fraction and significantly less MS identified protein species. Transiently and stably produced vector samples of MS set 1 and 2 were prepared by harvesting from a comparable number of producer cells and processing vector harvests for MS analysis in the same way including 40 fold concentration

by ultracentrifugation. As previously described, semi-quantitative comparison of viral proteins in each vector sample showed that p24 protein levels in stably produced samples were 6 fold lower compared to transiently produced samples.

Preparation of stably produced vector samples of MS set 3 included a six fold increase of concentration of vector harvests during ultracentrifugation to a final concentration factor of 240. This ensured that a comparable level of p24, thus a comparable number of vector particles in stably and transiently produced samples was subjected to LC-MS-MS and 10 µg of total protein per vector sample were analysed (MS set 3) (see chapter 3).

4.3.1.2. Sample Optimisation to reduce Salt Contents

Initially, exploratory work was carried out, in which SEC eluted void peak fractions were concentrated by diafiltration. Diafiltration is similar to dialysis applying a semi-permeable membrane to separate macromolecules from low molecular-weight compounds. Amicon Filter Units with a cut-off value of 3 kDa (Millipore) were used and vector samples concentrated by centrifugation. Unlike dialysis, Amicon Filter diafiltration does not rely on passive diffusion but on centrifugation, forcing water and molecules that are below 3 kDa through the membrane, resulting in concentration of large size molecules such as vector particles. This was followed by buffer exchange using dialysis with 50 mM Tris buffer of pH 8.5. Subsequent lyophilisation of dialysed samples reduced samples to the required volume of less than 200 µl. However salt levels were too high in these samples to allow sufficient trypsin digestion of sample proteins. Increasing the number of pooled void peak fractions to eight fractions per vector samples resulted in larger sample volumes and even higher amounts of salt per sample. Sample preparation using diafiltration was not further pursued. An alternative method of reducing the amount of salt in vector samples, dialysis, was carried out using ammonium bicarbonate (ABC) resulting in complete removal of salts and a pH of 8.0. Using larger dialysis cassettes made the concentration step by diafiltration between SEC and dialysis redundant.

4.3.2. LC-MS/MS Settings for Data Acquisition and Analysis

Mass spectrometry data analysis was carried out using Thermo Proteome Discoverer 1.2. with built-in Sequest (Thermo Fisher Scientific, Waltham, MA, USA). This program correlates tandem mass spectra of fragment ions with the predicted mass from amino acid sequences of peptides obtained from a database (Eng et al., 1994). For data analysis a database was composed including all known human proteins (taxon identifier 9606), bovine proteins (9913), proteins from HIV-1 group M subtype B (isolate HXB2) (11706), VS-Indiana virus (strain San Juan) (VSIV) (11285) and RD114 virus (11834), RDpro envelope protein (Ikeda et al., 2003) and Green fluorescent protein (GFP) (P42212) downloaded from Uniprot (Uniprot Consortium, 2012). Analysis and acquisition settings are described in chapter 2 (section 2.3.3.2.). A manufacture default filter setting of a minimum of 2 peptides per protein was applied for data acquisition requiring the detection of two different peptides per protein for confident protein identification. LC-MS/MS analysis of stably produced vector samples of MS set 1 resulted in identification of only 1 to 5 different cellular protein species. In comparison 14 to 49 different cellular protein species were detected in transiently produced samples. RDpro Env was not identified in stably produced vector samples when using the default MS settings. Hence for data analysis the initially used setting was relaxed so that a minimum of one peptide per protein was required to confidently identify a protein. However RDpro Env peptides were still not detected with the relaxed settings.

Another MS method of data acquisition was used in order to increase the likelihood to select less abundant proteins in a sample. As described in chapter 2 (section 2.3.3.2.), during the MS survey scan the five most abundant ions are chosen dynamically for fragmentation. A tandem MS spectrum of fragmented ions is created and its peptide sequence identified, hence only peptides whose ions were selected in the MS survey scan are identified. Ions that are not selected during a survey scan, for example ions of low abundance, will not be processed and corresponding peptides not be identified. The alternative MS method of data acquisition that

was used rejects ions that had been selected in one survey MS scan from re-selection in the next MS survey scan in favour of selecting ions of other peptides, increasing the likelihood of identifying proteins that are of lower abundance in the sample. This method (from now on referred to as 'rejection method') was only used in analysing RDpro purified vector samples in MS set 3 and allowed the identification of 4% and 7% of RDpro Env protein sequence in RDpro-GFP and RDpro-Empty vector samples, respectively (Table 13, column 'MS set 3 rejection').

4.3.3. LC-MS/MS Data Analysis

4.3.3.1. Identification of Viral Proteins in Purified LVs by LC-MS/MS

Peptides of viral proteins were detected in purified vector samples by LC-MSM/MS, including peptides of HIV-1 Gag protein domains matrix (MA), capsid (CA) and p2 as well as peptides of protease (PR), reverse transriptase (RT) and integrase (IN). Peptides of viral envelope proteins VSV-G Env and RDpro Env were also detected. In all samples, but VSV-G-only, peptides of HIV-1 Gag could be identified and was the most dominant protein apart from VSV-G-pseudotyped samples in which the VSV-G envelope protein was the most abundant protein detected. The percentage of covered protein sequence is shown in Table 13 for the samples of MS set 3 of transiently produced vectors (40 fold concentrated) and stably produced vectors (240 fold concentrated). The percentage of protein sequence that was identified using the rejection method for MS analysis of stably produced samples is also displayed (Table 13). Using the rejection method, more peptide sequences were identified covering a wider sequence range of a protein. For example, instead of 57% of peptide coverage, the peptide coverage of CA protein was increased to up to 79% using the rejection method. RDpro Env was only identified using the rejection method and peptides could be detected covering 4% to 7% of the RDpro amino acid (AA) sequence. Gag-p2 and PR were also only detected using the rejection method in stably produced samples. Coverage of CA protein sequence was the highest ranging between 57% and 79%, equivalent to 122 to 169 AA out of capsid's 215 AA. Peptides of Gag-p6 were never detected in stably produced samples but in transiently produced samples VSV-G-

GFP and VSV-G-Empty. As expected no peptides of VSV-G Env were detected in Gag/Pol-GFP samples but in all other transiently produced samples produced by co-transfection including VSV-G Env and no peptides originating from Gag or Gag-Pol protein were detected in the VSV-G-only sample.

Table 13: Coverage of Viral Protein Sequences by LC-MS/MS in Purified Vectors

viral protein		protein length (amino acids)	MS set 3		MS set 3 rejection		MS set 3			
			RDpro-GFP (240x)	RDpro-Empty (240x)	RDpro-GFP (240x)	RDpro-Empty (240x)	VSV-G-GFP (40x)	VSV-G-Empty (40x)	Gag/Pol-GFP (40x)	VSV-G-only (240x)
Gag	MA	130 AA	**45%**	**46%**	**63%**	**66%**	**73%**	**96%**	**67%**	ND
	CA	215 AA	**57%**	**57%**	**79%**	**61%**	**49%**	**50%**	**52%**	ND
	p2	13 AA	ND	ND	**100%**	**100%**	ND	ND	ND	ND
	p6	56 AA	ND	ND	ND	ND	**27%**	**18%**	ND	ND
Pol	PR	98 AA	ND	ND	**21%**	ND	**17%**	ND	**18%**	ND
	RT	559 AA	**3%**	**5%**	**33%**	**21%**	**17%**	**17%**	**17%**	ND
	IN	287 AA	ND	ND	**15%**	**9%**	**15%**	**9%**	**1%**	ND
VSV-G-Env		411 AA	ND	ND	ND	ND	**20%**	**22%**	ND	**41%**
RDpro Env		565 AA	ND	ND	**4%**	**7%**	ND	ND	ND	ND

Coverage of percentage of protein sequence is shown for vector samples of MS set 3; RDpro-pseudotypes are concentrated 240x and transiently produced vectors 40x by ultracentrifugation. The default MS analysis settings were used or the rejection method, as indicated. Detected viral proteins include (UniProt identifier): HIV-1 Gag-Pol polyprotein (P04585; NCBI Ref # NP_057849.4); VSV-G glycoprotein (P03522, NP_041715.1); RDpro protein sequence based on RD114 glycoprotein (A7LKA7) (Ikeda et al., 2003, Bell et al., 2010). ND= not detected.

4.3.3.2. Identified Host Cell Proteins

The results of MS detected cellular proteins are presented in Table 14. Mass spectrometry analysis did not allow quantitation of proteins within one sample, hence proteins were listed in the order of their detection in MS set 1 of vector samples and then grouped for example heat shock proteins listed together. For MS analysis each vector samples was subjected to one to three MS runs. Colour coding shows if a protein was detected in one (grey), two (yellow) or three runs (green) out of a total three MS runs of one sample. Of each transiently and stably produced vector sample three sets were made and analysed by MS (MS set 1 to 3). The VSV-G-only sample was analysed once. In total 93 different cellular proteins were identified for all samples and all sets regardless the number of runs they appeared in.

4.3.3.3. Variation in Host cell Protein Identification due to Sample Difference

Variation of results could be seen when analysing one vector sample. 1) Variations between MS runs of one set of a vector sample, for example Gag/Pol-GFP set 1 was analysed three times by MS. Ten cellular proteins were detected in all three MS runs (Table 14, green fields) and six cellular proteins in two out of three runs (Table 14, yellow fields). 2) Variations in the presence of a specific protein in different sets of one vector sample, for example VSVG-Empty. Many proteins were detected in set 1, 2 and 3 of VSV-G-Empty but some only in one set. For instance, 36 proteins were detected in one out of three sets and 26 proteins in more than one set. In VSV-G-GFP sets 35 proteins were detected only in one set and 18 proteins in more than one out of three sets. 3) Variations of the number of different host protein species between sets of a vector sample. The number of different detected host proteins varied in VSV-G-GFP sets between 21 and 38 different proteins and in VSV-G-Empty sets between 24 and 49 proteins.

Between different vector samples a certain level of similarity could be identified. VSV-G-GFP and VSV-G-Empty shared 39 cellular proteins. In Gag/Pol-GFP samples, that did not carry any envelope proteins, 10 or 14 cellular proteins were detected in MS set 1 and 3, respectively. Five cellular proteins were shared between these two sets. Since the viral protein VSV-G Env was detected in MS set 2 of Gag/Pol-GFP, albeit at relatively low abundance, the results of this sample are excluded from further analysis. Only myosin-9 was unique to Gag/Pol-GFP vector sets. However all of the other 19 different cellular protein species detected in Gag/Pol-GFP set 1 and 2 were also detected in VSV-G-GFP and/or VSV-G-Empty. One set of VSV-G-only sample was prepared, concentrated 240 fold and analysed by MS. VSV-G Env was the most frequently detected protein. In total 61 different cellular protein species were identified in this sample. VSV-G-GFP, VSV-G-Empty and VSV-G-only shared 38 proteins. Nineteen of these proteins were not detected in the samples Gag/Pol-GFP and RDpro-pseudotypes, thus only detected in all three VSV-G pseudotypes. These included proteins such as Actin-alpha, Cofilin 2, Calmodulin 1, Neutral amino acid transporter, Solute carrier family 3, Glyceraldehyde-3-phosphate dehydrogenase and Triosephosphate isomerase. Interestingly, Neutral amino acid transporter, also known as RD114/simian type D retrovirus receptor and the heavy chain of Solute carrier family 3 have been only detected in VSV-G pseudotyped samples. Thirteen proteins were unique to the 240 fold VSV-G-only sample.

All proteins that were identified in stably produced samples were also identified in transiently produced samples with the exception of AHNAK (or Desmoyokin; Neuroblast differentiation-associated protein) a protein only identified in RDpro-GFP and RDpro-Empty samples of set 3 (240 fold concentrated). Peptides of RDpro Env were only detected when using the rejection method and after relaxing the applied filters by lowering the minimum number of identified peptides per protein for its unambiguous identification from two to one. Consequently one peptide of the RDpro protein could be detected at high confidence or one or more peptides at low confidence in MS set 3 of RDpro-GFP and RDpro-Empty.

Finally some identified cellular proteins were common to all five stably and transiently produced vector samples containing HIV-1 structural proteins (VSV-G-GFP, VSV-G-Empty, Gag/Pol-GFP and RDpro-GFP and Empty)**,** regardless of the concentration factor or MS detection method. These include Beta-Actin, Heat shock cognate 71, Heat shock protein 70, Histone cluster 1- H2ah, Clathrin heavy chain 1, Alpha-enolase (ENO1), ALIX, Elongation factor 1-alpha (EEF1A) and Cyclophilin A (Cyp A). MARCKSL1 was common to VSV-G-GFP, VSV-G-Empty as well as RDpro-GFP and RDpro-Empty LV samples. Several bovine proteins have been detected in vector samples for example Serum albumin, Hemoglobin fetal subunit alpha and beta, potentially carried over from fetal calf serum used in cell culture during vector production and are excluded from Table 14. Similar numbers of different protein species were detected in VSV-G and RDpro- pseudotypes when using the rejection method for RD-pro sample analysis (see section 4.3.2.) Overall, the same bovine protein species have been detected in VSV-G-pseudotypes and RDpro-pseudotypes. No cellular proteins were identified that are unique to vector samples that carry a vector genome (GFP transfer vector), suggesting that none of the detected cellular proteins is solely associated with viral RNA.

4.3.3.4. Difference in Host cell Protein Identification due to Technical Variation

The second type of variation in MS results of vector samples can be classed as technical variation. This includes 1) preparation of vector sample with a concentration factor of 40 fold or 240 fold and 2) use of default settings or rejection method for MS analysis. MS analysis of 40 fold concentrated stably produced samples detected a low number of cellular protein species varying from 1 to 5. On average 4.5 µg total protein were analysed in MS set 1 and 2. In order to exclude the possibility that this was due to insufficient amounts of viral particles in the prepared sample, RDpro-pseudotyped vector concentration was increased from 40 to 240 fold during production. Concentration of RDpro-pseudotyped vectors was increased by 6 fold to a final concentration factor of 240 by ultracentrifugation as previously described (see

chapter 3) and therefore 10 µg of total protein were subjected to MS (MS set 3). Compared to results of 40 fold concentrated RDpro pseudotyped samples, set 1 and set 2, a six fold increased concentration factor resulted in a small increase in the number of protein species detectable by LC-MS/MS. In 240 fold concentrated RDpro-GFP and RDpro-Empty samples 5 and 8 cellular proteins species were identified, respectively, however compared to transiently produced samples this number is still substantially lower. This could indicate that RDpro-pseudotyped vectors associate with less cellular proteins than VSV-G pseudotypes, however this hypothesis could only be confirmed if an additional sample, RDpro-GFP produced by transient transfection, was included in this study. To investigate if RDpro-pseudotypes contain many protein species that are less frequently detected, hence potentially of low abundance, MS set 3 of 240 fold stably produced samples was analysed using the above described rejection method of data collection. This increased the number of detected viral and cellular peptides to 15 different cellular protein species detected in RDpro-GFP and 19 cellular proteins in RDpro-Empty, comparable to Gag/Pol-GFP vectors.

Table 14: LC-MS/MS Identified Proteins in all Vector Samples (listed in the order of their detection in MS set 1 of vector samples, then grouped)

total number of cellular/host proteins		21	23	38	49	24	27	61	14	10	1	4	5	15	5	2	8	19
		VSV-G-GFP			VSV-G-Empty			VSV-G - only	Gag-Pol-GFP		RDpro-GFP				RDpro-Empty			
Full Protein Name	**UniProt ID**	**set 1**	**set 2**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3**	**set 1**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3 rejection**	**set 1**	**set 2**	**set 3**	**set 3 rejection**
Glycoprotein G (VSV-G)	P03522																	
Gag-Pol polyprotein (HIV-1)	P04585																	
RDpro envelope (RD114)	A7LKA7																	
Actin, beta (ACTB)	Q53G76																	
Actin, alpha skeletal muscle (ACTA1)	Q5T8M8																	
Cofilin 1 , non-muscle (CFL1)	G3V1A4																	
Cofilin 2, muscle (CFL2)	G3V5P4																	
Calmodulin 1 (CALM1)	E7ETZ0																	
ATPase, Na+/K+ transporting, alpha 1 polypeptide (ATP1A1)	B7Z3U6																	
ATPase, Na+/K+ transporting, beta 3 polypeptide (ATP1B3)	C9JA36																	
Heat shock cognate 71kDa protein (HSC71)	P11142																	
Heat shock 70 kDa protein 1A/1B(HSPA1A)	P08107																	
Heat shock protein HSP 90-beta (HSP90AB1)	P08238																	
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ)	E7EX29																	
Triosephosphate isomerase (TPI1)	B7Z5D8																	

Table 14 continued (2/5):

Full name	UniProt ID	V+ set 1	V+ set 2	V+ set 3	V- set 1	V- set 2	V- set 3	V set 3	G/P set 1	G/P set 3	R+ set 1	R+ set 2	R+ set 3	R+ set 3 R	R- set 1	R- set 2	R- set 3	R- set 3 R
Histone cluster 1, H1c (HIST1H1C)	P16403																	
H2A histone family, member X (H2AFX)	P16104																	
Histone cluster 1, H2ah (HIST1H2AH)	Q96KK5																	
Histone H2B (HIST2H2BE)	A8K9J7																	
Histone cluster 1, H2bj (HIST1H2BJ)	P06899																	
Histone cluster 1, H4h (HIST1H4H)	Q0VAS5																	
Neutral amino acid transporter	Q15758																	
Solute carrier family 3 (AA transporter heavy chain), member 2 (SLC1A5)	F5GZS6																	
60S acidic ribosomal protein P2 (RPL2)	P05387																	
60S ribosomal protein L7a (RPL7A)	P62424																	
60S ribosomal protein L3 (PRL3)	P39023																	
Ribosomal protein L4 (RPL4)	P36578																	
Ribosomal protein L18 (PRL18)	F8VYV2																	
Ribosomal protein S8/S3 (RPS8/S3)	Q5JR95																	
Ribosomal protein L14 (RPL14)	Q6IPH7																	
Nucleophosmin (NPM1)	P06748																	
Elongation factor 1-alpha (EEF1A)	Q504Z0																	
Transferrin receptor protein 1 (TFRC)	P02786																	
Polyubiquitin-C (UBC)	P0CG48																	
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	E7EUT5																	
Tubulin, alpha 1b (TUBA1B)	Q8WU19																	
Tubulin, alpha 1a (TUBA1A)	Q71U36																	

Table 14 continued (3/5):		V+			V-			V	G/P		R+				R-			
Full name	**UniProt ID**	**set 1**	**set 2**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3**	**set 1**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3 R**	**set 1**	**set 2**	**set 3**	**set 3 R**
Tubulin, beta class I (TUBB)	Q5JP53																	
Syntenin-1 (SDCBP)	B4DHN5																	
Chloride intracellular channel protein 1 (CLIC1)	O00299																	
Alpha-enolase (ENO1)	P06733																	
Guanine nucleotide-binding protein G(k) subunit alpha (GNAI3)	P08754																	
Clathrin heavy chain 1 (CLTC)	Q00610																	
Programmed cell death 6-interactingprotein (ALIX)	Q8WUM4																	
Neuroblast differentiation-associated protein (AHNAK)	Q09666																	
Radixin (RDX)	P35241																	
MARCKS	P29966																	
MARCKS-related protein (MARCKSL1)	P49006																	
Annexin A2 (ANXA2)	P07355																	
Annexin A5 (ANXA5)	E9PHT9																	
Cyclophilin A (CYPA)	A8K486																	
Moesin (MSN)	P26038																	
Basigin (BSG)	P35613																	
Ezrin (EZR)	E7EQR4																	
Ras-related protein Rab-1A (RAB1A)	B7Z8M7																	
Ras-related protein Ral-A (RALA)	P11233																	
Ras-related protein R-Ras2 (RRAS2)	B7Z6C4																	
Ras-related protein Rab-7a (RAB7A)	P51149																	
Cortactin (CTTN)	Q96H99																	
Y box binding protein 1 (YBX1)	P67809																	

Table 14 continued (4/5):		V+			V-			V	G/P		R+				R-			
Full name	**UniProt ID**	**set 1**	**set 2**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3**	**set 1**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3 R**	**set 1**	**set 2**	**set 3**	**set 3 R**
Brain abundant, membrane attached signal protein 1 (BASP1)	P80723																	
unconventional Myosin IC (MYO1C)	F5H6E2																	
Myosin-9 (MYH9)	P35579																	
Desmoglein-2 (DSG2)	Q14126																	
Poly [ADP-ribose] polymerase 1 (PARP1)	P09874																	
Protein kinase C and casein kinase substrate in neurons 3 (PACSIN3)	Q9UKS6																	
Vesicle-associated membrane protein 3 (VAMP3)	Q15836																	
Plastin 3 (PFS3)	B7Z6M1																	
Profilin-1 (PFN1)	P07737																	
Prostaglandin F2 receptor inhibitor (PTGFRN)	Q4QQP8																	
Creatine kinase B-type (CKB)	P12277																	
Filamin A (FLNA)	Q5HY54																	
Lactadherin (MFGE8)	Q08431																	
Charged multivesicular body protein 4b (CHMP4B)	Q9H444																	
EH domain-containing protein 4 (EDH4)	Q9H223																	
Zinc finger CCCH-type antiviral protein 1 (ZC3HAV1)	Q7Z2W4																	
Mov10, Moloney leukaemia virus 10, homolog (mouse) (MOV10)	Q9HCE1																	
Talin 1 (TLN1)	Q5TCU6																	
Polyadenylate-binding protein 4 (PABPC1)	P11940																	
Ubiquitin-like modifier activating enzyme 1	Q5JRR6																	
Tubulin, beta 3 class III (TUBB3)	Q9BV28																	

Table 14 continued (5/5):		V+			V-			V	G/P		R+				R-			
Full name	**UniProt ID**	**set 1**	**set 2**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3**	**set 1**	**set 3**	**set 1**	**set 2**	**set 3**	**set 3 R**	**set 1**	**set 2**	**set 3**	**set 3 R**
Plasma membrane calcium-transporting ATPase 1 (ATP2B1)	P20020																	
Golgi-associated, gamma adaptin ear containing, ARF binding protein 2 (GGA2)	O14564																	
Secretory carrier-associated membrane protein 3 (SCAMP3)	O14828																	
Ras-related C3 botulinum toxin substrate 1 (RAC1)	A4D2P2																	
Ras-related protein Rap-1b (RAP1B)	B4DW94																	
Alpha-actinin-4 (ACTN4)	A4K467																	
Lipolysis stimulated lipoprotein receptor (LSR)	Q86X29																	
Destrin (actin depolymerizing factor) (DSTN)	F6RFD5																	
Erythrocyte band 7 integral membrane protein	P27105																	
Fructose-bisphosphate aldolase A (ALDOA)	P04075																	
Importin subunit alpha-1 (KPNA2)	Q53YE3																	
Solute carrier family 9, subfamily A (SLC9A3R1)	O14745																	
Sorting nexin-9 (SNX9)	Q9Y5X1																	
Transgelin-2 (TAGLN2)	P37802																	
WD repeat-containing protein 1 (WDR1)	O75083																	

LC-MS/MS identified proteins listed in the order of their detection in MS set 1 of vector samples, then grouped into for example heat shock proteins, ribosomal proteins. UniProt accession number for each protein is given (http://www.uniprot.org/). Between 1 and 3 preparations per vector sample were analysed (MS sets). Identified proteins in RDpro-GFP and RDpro-Empty samples MS set 3 using the rejection method are indicated ('set 3R'). Colour coding: a protein was detected in one (grey), two (yellow) or three runs (green) out of a total three runs of one set. Striped shading of RDpro Env in MS set 3R indicates only one peptide of this protein was detected at high confidence or 2 or more at low confidence.

In total, five cellular proteins, AHNAK, ALIX, MARCKSL1, ENO1 and EEF1A, were selected for further analysis that were common to most of the analysed samples and have not been associated with the HIV-1 life-cycle, in particular with assembly or budding of HIV-1 before. However they have been previously identified in HIV-1 virions, thus were interesting for further analysis. AHNAK, unique to RDpro-pseudotyped samples, was also further analysed. The coverage of their protein sequence by identified peptides is shown in Table 15. Peptides covering 1.8 % to 3.4 % of the 5890 AA long AHNAK sequence were detected in RDpro-pseudotypes and none in samples with VSV-G Env or Gag-Pol/GFP without an envelope protein; hence this protein may be associated with RDpro Env. Using the default MS acquisition method ALIX, MARCKSL1, ENO1 and EEF1A were identified in at least 4 out of the six different analysed vector samples. Applying the rejection methods for MS detection ALIX, ENO1 and EEF1A could be also detected in stably produced samples.

Table 15: Coverage of Selected Host Protein Sequences (in %) detected by LC-MS/MS in Purified Vectors

protein	protein length	MS set 1*, 2** or 3				MS set 3		MS set 3 rejection	
		VSV-G – GFP 40x	VSV-G – Empty 40x	Gag/Pol-GFP 40x	VSV-G – only 240x	RDpro-GFP 240x	RDpro-Empty 240x	RDpro – GFP 240x	RDpro - Empty 240x
AHNAK	5890 AA	ND	ND	ND	ND	1.8%	3.4%	1.8%	1.2%
ALIX	873 AA	5.2%	6.7%	6.0%	3.7%	ND	ND	4.7%	6.5%
MARCKSL1	195 AA	39%	11%*	ND	47%	37%	45%	36%	44%
ENO1	434 AA	6.9%	7.3%	8.8%	13%	ND	ND	12%	14%
EEF1A	308 AA	7.6%**	11%	7.6%*	11%	ND	ND	11%	11%

Proteins detected in LC-MS/MS of each MS set vector sample: in 3 of 3 MS runs of each sample (green); 2 of 3 runs (orange); 1 of 3 runs (grey); protein length is given in amino acids (AA). Protein detected in set 1 (*) or set 2 (**).

4.3.5.3. Confirmation of the Presence of Selected Host Proteins

Western blotting of p24 and envelope proteins VSV-G as well as RDpro in vector samples was used to verify the presence of viral proteins (see chapter 3). The presence of selected MS-identified host-proteins, AHNAK and ALIX, was also confirmed by Western blotting (Figure 17).

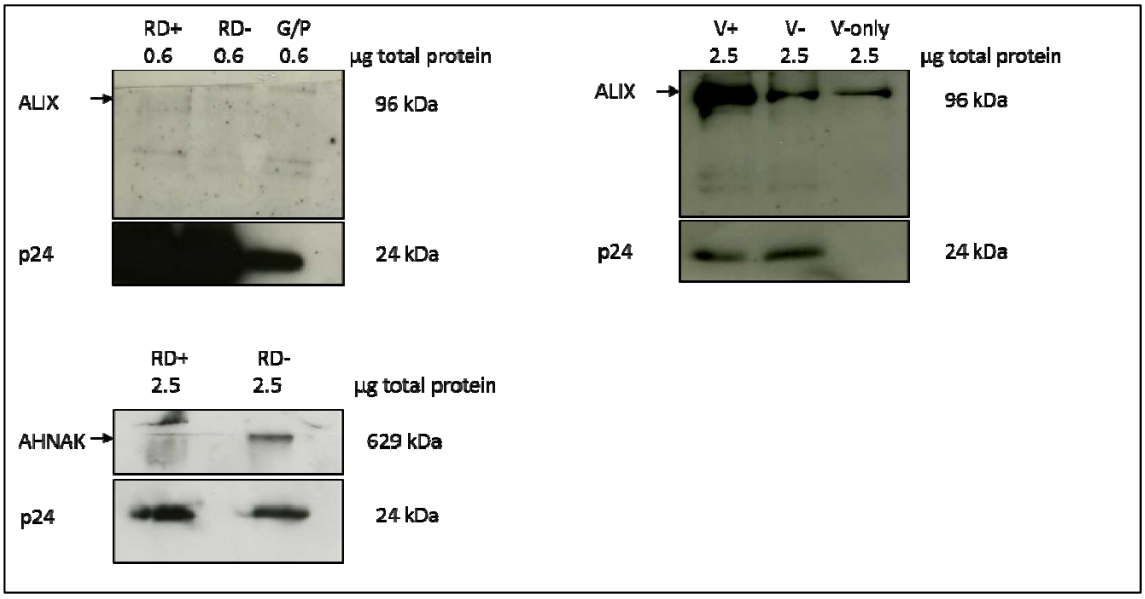


Figure 19: Confirmation of LC-MSM/MS detected Host Proteins by Western Blotting in vector samples. ALIX was detected in purified vector samples by Western blotting in stably and transiently produced samples. The presence of AHNAK was confirmed in stably produced purified vector samples. RDpro-GFP (RD+), RDpro-Empty (RD-), Gag-Pol-GFP (G/P), VSV-G-GFP (V+), VSV-G-Empty (V-), VSV-G-only (V-only).

4.3.5.4. Functions of Identified Host Proteins

For functional characterisation and analysis of LC-MS/MS identified proteins, IPA ('Interactive Pathway analysis' of complex 'omics data by Ingenuity® Systems; www.ingenuity.com; Mountain View, CA, USA) was used. Protein function analysis is based on functional annotations in the web-based Ingenuity® Knowledge Database. All identified

cellular proteins were collated in a list using the UniProt identifier and uploaded into IPA. The 93 identified cellular proteins were mapped by IPA, of which 52 are located in the cytoplasm, 15 in the nucleus, 24 at the plasma membrane and two proteins were classified as being located in the extracellular space (Lactadherin and WD repeat-containing protein 1). Identification data was further categorised into functional families. Out of 93 proteins, 23 were classified as enzymes, 12 as transporters, 5 as transcription regulators and the remaining 53 proteins were not classified.

Cellular functions were identified and categorised. Categories with the most proteins of our dataset are shown in Table 16. About 51% LC-MS/MS identified proteins take part in the mechanisms of cellular assembly and organisation, cell function and maintenance as well as cell death. Some of the proteins fall into several functional categories as shown in Figure 18, hence identified host proteins have more than one specific role. These identified proteins are annotated to function in processes such as the organisation of the cytoplasm and cytoskeleton, in microtubule dynamics, formation of cellular protrusions and in necrosis and apoptosis.

Table 16 : Top Molecular and Cellular Functions of LC-MS/MS Identified Proteins

Functional Category	Functions Annotation		No. of Proteins
Cellular Assembly and Organisation	Organisation of Cytoplasm, Organisation of Cytoskeleton, Microtubule Dynamics, Formation of Cellular Protrusions	Organisation of Organelle	45
Cellular Function and Maintenance		Engulfment of Cell, Phagocytosis, Endocytosis, Organisation of Filaments	51
Cell Death	Necrosis, Apoptosis		50

Top functional categories with at least 10 proteins per function annotation identified by LC-MS/MS are shown using Ingenuity® Knowledge Database of IPA by Ingenuity® Systems.

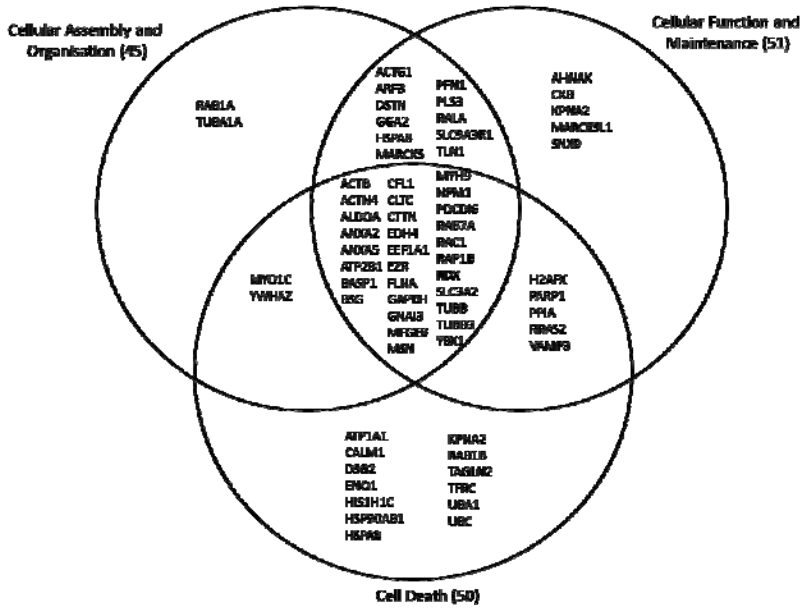


Figure 20: Venn diagram of top molecular and cellular functions of LC-MS/MS identified proteins classified by functional annotation of the Ingenuity® Knowledge Database of IPA by Ingenuity® Systems. Functional overlap of proteins is shown.

The IPA software maps the proteins in the dataset to the information in the Ingenuity®

Knowledge Base and then places the identified cellular proteins into well-established signalling or metabolic pathways, termed "canonical pathways". Canonical pathways have been defined as “idealised or generalised pathways that represent common properties of a particular signalling module or pathway” (Science magazine (http://stke.sciencemag.org/about/help/cm.dtl) as opposed to specific pathways in which components are known to act together in a particular organism, tissue or cell type.

The top five pathways each with at least 10 associated proteins of the dataset are displayed along with the number of proteins from the dataset that map to each pathway over the total number of proteins that map to the canonical pathway in the Ingenuity® Pathways Knowledge Base is displayed in Figure 19A. These pathways include actin cytoskeleton signalling, epithelial adherence junction signalling and remodelling (Figure 19B), integrin signalling and clathrin mediated endocytosis (Figure 19C). Proteins associated with mechanisms of viral host cell exit and their association with the top five pathways is illustrated. Several proteins act in more than one pathway as depicted in the Venn diagrams. In this study proteins have been identified that have previously been shown to be involved in viral exit from host. Among them are two known factors involved in the ESCRT machinery, ALIX, identified in all vector samples as well as the VSV-G-only control and CHMP4, member of ESCRT-III, only identified in VSV-G-GFP and VSV-G-Empty samples. Members of the actin protein family, actin alpha and actin beta, have been detected in all vector samples.

A

Canonical Pathway	**Identified Proteins/ Total Number of Proteins of Canonical Pathway in** Ingenuity® Pathways Knowledge Base
Actin Cytoskeleton Signalling	15/242
Epithelial Adherent Junction Signalling	12/154
Epithelial Adherent Junction Remodelling	10/70
Clathrin Mediated Endocytosis	11/198
Integrin Signalling	11/208
Mechanism of Viral Exit from Host Cell	4/45

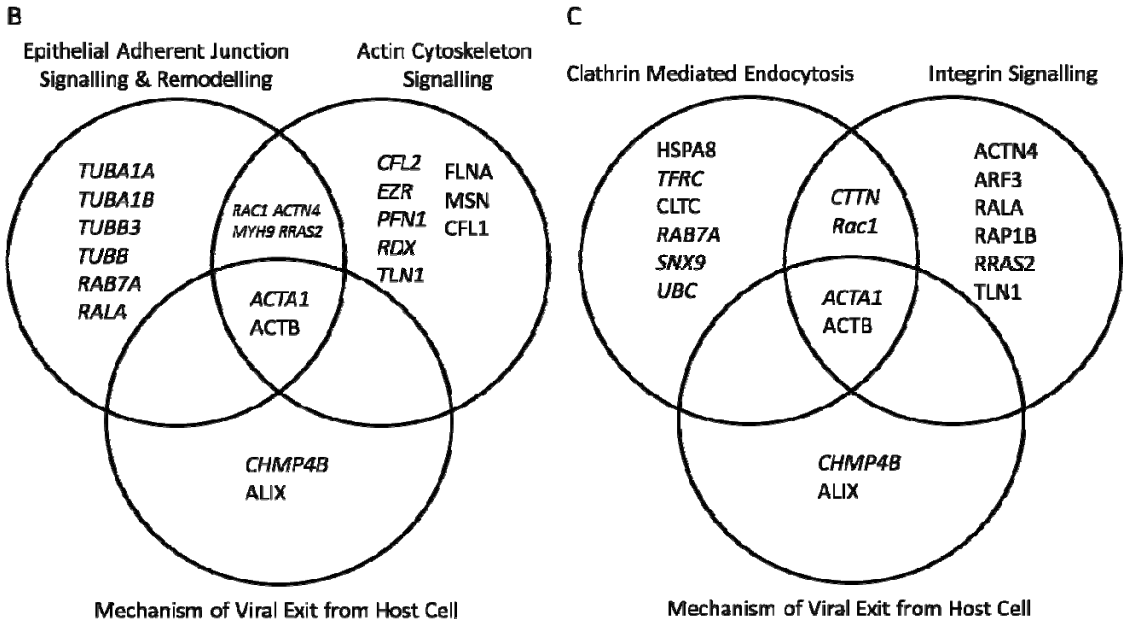


Figure 21: LC-MS/MS Identified Proteins of the Top Five Canonical Pathways. A) Top five canonical pathways, at least 10 of the identified host proteins are associated to the listed pathways. Furthermore four proteins were identified out of 45 proteins that are associated with mechanisms of viral exit from host cells stored in the Ingenuity® Pathways Knowledge Base. B) Venn diagrams of proteins that are associated with at least one of the top five canonical pathways.

4.3.5.5. Functions of Proteins Exclusive to Samples Containing VSV-G (VSV-G-GFP, VSV-G-Empty and VSV-G-only)

Functional category and pathway analysis of 19 proteins, that were exclusively identified in all three VSV-G Env-containing samples, including VSV-G-GFP, VSV-G-Empty and VSV-G-only, was performed using the Ingenuity® Pathway Knowledge Base. Ten out of 19 of these proteins are located at the plasma membrane, eight proteins in the cytoplasm and one protein in the nucleus. Functional categorisation using IPA software assigns these 19 proteins to be involved in cellular assembly and organisation, such as organisation of the cytoskeleton, microtubules dynamics and formation of cellular protrusions but also molecular transport (Table 17). Actin-alpha, Cofilin 2, Radixin, Ezrin, Tubulin-beta class I, Protein kinase C and casein kinase substrate and Talin 1 are members of the cytoskeleton or take part in its organisation. Four of the proteins found in VSV-G containing samples work as transporters, including transmembrane transporters (SLC1A5, SLC3A2 and Basigin) as well as proteins helping vesicular transport, such as Vamp3.

Table 17: LC-MS/MS Identified Proteins Unique to Samples Containing VSV-G

Full Name	UniProt Identifier	Family	Cellular Location	Function (NCBI Gene Annotation)
Actin, alpha skeletal muscle (ACTA1)	Q5T8M8	Other	Cytoplasm	structural constituent of cytoskeleton
Cofilin 2, muscle (CFL2)	G3V5P4	Other	Cytoplasm/ Nucleus	involved in the regulation of actin-filament dynamics, controls actin polymerization and depolymerisation in a pH-dependent manner
Calmodulin 1 (CALM1)	E7ETZ0	Other	Cytoplasm	member of the EF-hand calcium-binding protein family
Triosephosphate isomerase (TPI1)	B7Z5D8	Enzyme	Cytoplasm	catalyses the isomerization of glyceraldehydes 3-phosphate (G3P) and dihydroxy-acetone phosphate (DHAP) in glycolysis and gluconeogenesis
Neutral amino acid transporter (SLC1A5)	Q15758	Trans-porter	Plasma Membrane	neutral amino acid transporter, can act as a receptor for RD114/type D retrovirus (Larriba et al., 2001)
Solute carrier family 3 (amino acid transporter heavy chain), member 2 (SLC3A2)	F5GZS6	Trans-porter	Plasma Membrane	regulation of intracellular calcium levels and transports L-type amino acids
Ras-related protein Rab-1A (RAB1A)	B7Z8M7	Enzyme	Cytoplasm	member of the Ras superfamily of GTPases, cycles between inactive GDP-bound and active GTP-bound forms, vesicle traffic from endoplasmic reticulum to Golgi apparatus
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	E7EUT5	Enzyme	Cytoplasm	catalyses energy-yielding step in carbohydrate metabolism, reversible oxidative phosphorylation of glyceraldehyde-3-phosphate in the presence of phosphate and nicotinamide adenine dinucleotide (NAD)
Tubulin, beta class 1 (TUBB)	Q5JP53	Other	Plasma Membrane	major components of microtubules
Radixin (RDX)	P35241	Other	Cytoplasm	may be important in linking actin to plasma membrane, highly similar in sequence to both ezrin and moesin

Table 17 continued (2/2):

Full Name	UniProt Identifier	Family	Cellular Location	Function (NCBI Gene Annotation)
MARCKS	P29966	Other	Plasma Membrane	substrate for protein kinase C, actin filament crosslinking protein
Basigin (BSG)	P35613	Trans-porter	Plasma Membrane	important in spermatogenesis, embryo implantation, neural network formation, tumour progression
Ezrin (EZR)	E7EQR4	Other	Plasma Membrane	cross-linkers between plasma membranes and actin-based cytoskeletons, key role in cell surface structure adhesion, migration and organisation
Desmoglein-2 (DSG2)	Q14126	Other	Plasma Membrane	part of desmosomes, cell-cell junctions between epithelial cells
Poly [ADP-ribose] polymerase 1 (PARP1)	P09874	Enzyme	Nucleus	important for differentiation, proliferation and tumour transformation, part of DNA strand break repair pathway
Protein kinase C and casein kinase substrate in neurons 3 (PACSIN3)	Q9UKS6	Other	Cytoplasm	involved in linking actin cytoskeleton with vesicle formation
Vesicle-associated membrane protein 3 (VAMP3)	Q15836	Other	Plasma Membrane	member of SNARE (Soluble NSF Attachment Protein Receptor), involved in vesicular transport from late endosomes to trans-Golgi network
EH domain-containing protein 4 (EHD4)	Q9H223	Enzyme	Plasma Membrane	role in early endosomal transport
Talin 1 (TLN1)	Q5TCU6	Other	Plasma Membrane	cytoskeletal protein, concentrated in areas of cell-substratum and cell-cell contacts, plays a significant role in actin filament assembly and in spreading and migration of various cell types

Proteins identified in all three samples containing VSV-G with UniProt identifier and cellular location are shown; proteins were detected in at least one MS sample, functional annotation taken from NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/gene/) or other sources as indicated.

4.3.3.5. Functions of Host Proteins Common to Transiently and Stably Produced LV Samples

Eight cellular proteins are common to all transiently and stably produced vector samples as well as VSV-G-only control including Beta actin, HSC71, HSPA1A, HIST1H2AH, Alpha-enolase, ALIX and Cyclophilin A. Clathrin heavy chain 1 was detected in all vector samples but VSV-G-only control. Elongation factor 1 alpha and MARCKSL1 were common to VSV-G-GFP, VSV-G-Empty as well as RDpro-GFP and RDpro-Empty LV samples and their individual functions are shown in Table 18. Categorising these proteins using IPA software shows that six of them are functioning in cellular proliferation. HSPA1A and HIST2AHA are not categorised by IPA but with chaperone activity and DNA structure organisational functions, respectively, they are also essential for cell maintenance. Three chaperones proteins HSC71, HSPA1A and CypA are common to all LC-MS/MS analysed LV samples. Beta actin and ALIX are involved in viral exit from cells. ALIX is a factor of the Endosomal Sorting Complexes Required for Transport (ESCRT) and functions in multivesicular body (MVB) biogenesis.

Table 18: Functions of Proteins Common to Transiently and Stably Produced LV Samples

Full Name	UniProt Identifier	Family	Cellular Location	Function (NCBI Gene Annotation)
Actin, beta (ACTB)	Q53G76	Other	Cytoplasm	member of actin protein family, plays a role in cell motility, structure and integrity
Heat shock cognate 71 kDa protein (HSC71)	P11142	Enzyme	Cytoplasm	chaperone, ATPase in the disassembly of clathrin-coated vesicles during transport of membrane components through the cell
Heat shock 70 kDa protein 1A/1B (HSPA1A)	P08107	Enzyme	Cytoplasm	stabilizes existing proteins against aggregation, mediates folding of newly translated proteins in the cytoplasm and in organelles, involved in the ubiquitin-proteasome pathway
Elongation factor 1-alpha (EEF1A)	Q504Z0	Trans-lation regulator	Cytoplasm	isoform of the alpha subunit of the elongation factor-1 complex, which is responsible for the enzymatic delivery of aminoacyl tRNAs to the ribosome
Histone H2ah (HIST2AHA)	Q96KK5	Other	Nucleus	basic nuclear proteins that are responsible for the nucleosome structure of the chromosomal fiber in eukaryotes
Clathrin heavy chain 1 (CLTC)	Q00610	Other	Plasma Membrane	major protein component of cytoplasmic face of intracellular organelles (coated vesicles and coated pits), intracellular trafficking of receptors and endocytosis of macromolecules
Alpha-enolase (ENO1)	P06733	Enzyme	Cytoplasm	homodimer composed of two alpha, two gamma or two beta subunits, functions as a glycolytic enzyme
Programmed cell death 6-interacting protein (ALIX)	Q8WUM4	Other	Cytoplasm	functions within ESCRT pathway, in intraluminal endosomal vesicle formation and in enveloped virus budding
MARCKS-related protein (MARCKSL1)	P49006	Other	Cytoplasm	member of myristoylated alanine-rich C-kinase substrate (MARCKS) family, cytoskeletal regulation, protein kinase C signalling and calmodulin signalling, formation of adherent junction
Cyclophilin A (CypA)	A8K486	Enzyme	Cytoplasm	member of the peptidyl-prolyl cis-trans isomerase (PPIase) family, catalyse the cis-trans isomerization of proline imidic peptide bonds in oligopeptides, accelerate protein folding

Proteins identified in all five VSVG-G as well as RDpro-pseudotyped vector samples with UniProt identifier and cellular location; proteins were detected in at least one replicate of each vector sample, functional annotation taken from NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/gene/).

4.3.3.6. AHNAK - Cellular Protein Unique to RDpro-VSV-G and RDpro-Empty

AHNAK or Neuroblast differentiation-associated protein has been identified only in stably produced LV samples pseudotypes with RDpro Env. Analysis of the IPA database shows that it has been found in several subcellular locations such as the nucleus, cytoplasmic vesicles and intercellular junctions.

4.4. Discussion

The aim of this chapter was to analyse the LC-MS/MS results comprising the cellular protein composition of six different lentiviral vectors purified by size exclusion chromatography (SEC). Any differences or similarities in host cell protein composition were investigated in an attempt to use this data as a basis for further analysis of their function in vector assembly and budding from the producer cell.

Quality of Results. Initially this study aimed at determining differences between vector production systems, between transiently and stably produced vectors, and pseudotypes. Many cellular proteins were repeatedly identified in the same vector sample among different MS sets or production batches of one vector sample and were also shared between different vector samples indicating their stable association with these vector samples. VSV-G-GFP, VSV-G-Empty and VSV-G-only had 38 cellular proteins in common. Variations among different vector samples could be due to differences in the production system or vector components (VSV-G or RDpro Env) but also due to different expression levels of some of the cellular proteins in the STAR and 293T producer cells. STAR cells have been modified to contain several additional expression cassettes that may affect cell characteristics resulting in the differential expression of some proteins compared to 293T cells. However, the observed variations in the identified proteins between different MS sets of one sample indicate that analysis of three batches of a vector may not be sufficient to determine which proteins are truly unique for each vector.

The type of MS analysis used here does not allow an absolute quantification of proteins in each sample, however viral proteins were the most frequently detected proteins in all vector samples confirming that viral vectors are major components of MS analysed samples. Gag and VSV-G Env were identified in the corresponding samples as expected, that is peptides of Gag-Pol polyprotein were detected in all samples apart from VSV-G-only. VSV-G Env was detected in the samples VSV-G-GFP, VSV-G-Empty and VSV-G-only but not in Gag/Pol-GFP.

Most dominant functional categories of identified proteins. This comparative analysis shows that the identified host cell proteins function predominantly in organisation of cytoskeleton and are involved in actin cytoskeleton signalling, including members of the actin and tubulin family as well as actin interacting proteins. Tubulins are a major component of the microtubules, part of the cytoskeleton. Proteins, vesicles or cell organelles are transported along microtubules and actin filaments towards the plasma membrane (Welte, 2004). Interaction with tubulin and viral proteins during transport to the assembly site (Jolly et al., 2007) could cause incorporation of tubulins into vectors. Also many of the identified proteins, such as ALIX, Rab1A, Rab7A, unconventional myosin IC, Vamp3 and GGA2, Clathrin heavy chain 1are found in the plasma membrane or function in vesicle trafficking.

MS analysis identified certain proteins that are particularly interesting for further analysis of their function or potential role in LV assembly. Proteins that were detected in most HIV-1 Gag containing LV samples and have been less well studied regarding HIV-1 assembly include Alpha-enolase, ALIX and Elongation factor 1-alpha. MARCKSL1 was common to all enveloped samples, regardless if pseudotyped with VSV-G or RDpro envelopes.

Literature analysis of the MS identified proteins in HIV-1 assembly. The involvement of some of the here identified proteins in HIV-1 assembly has been documented. Incorporation of actin into HIV-1 particles and binding of actin by Gag was documented (Wilk et al., 1999). Functional significance of the host cell cytoskeleton in HIV-1 assembly was then confirmed by fixed cell immunofluorescence labelling and confocal microscopy (Jolly et al., 2007). After inhibition of actin and tubulin in HIV-1 infected T cells assembly of Gag and Env at lipid raft-like domains at the plasma membrane was disrupted. Depolymerisation of actin decreased Gag release and virion infectivity. Actin and tubulin inhibitors reduced Env incorporation, suggesting that HIV-1 follows microtubule-directed routes within the cytoplasm towards the plasma membrane (Jolly et al., 2007). Chaperones HSP70 and HSC71, along other heat shock proteins, have been shown by western blotting to be incorporated into virion membranes and might bind HIV-1

Gag, holding it in assembly-competent conformation during transport to plasma membrane (Gurer et al., 2002). Western blotting has also confirmed the binding of Cyclophilin A to HIV-1 Gag of sucrose purified HIV-1 virions produced by HeLa cells (Franke et al., 1994). In contrast, other cytoplasmic abundant prolyl isomerases were not incorporated suggesting a specific interaction with HIV-1 Gag. CypA incorporates into virions by binding to the MA/CA domain of Pr55Gag (Luban et al., 1993), however the role of MA/CA-bound CypA during assembly has not be determined. In early steps of the HIV-1 life cycle CypA binds CA of HIV-1 Gag post-infection directing the nuclear import of the HIV-1 preintegration complex (PIC) (Schaller et al., 2011a) but this interaction is believed to involve CypA from the infected cell rather than the virion producer cell. EEF1A was shown to bind MA and NC of HIV-1 Gag (Cimarelli and Luban, 1999) and analysis of effects on replication of HIV-1 Gag mutants that cannot bind EEF1A showed reduced levels of viral proteins in cell supernatants. The exact role of this interaction in viral replication could not be determined yet. ALIX, a factor of the Endosomal Sorting Complexes Required for Transport (ESCRT) is known to assist in the viral budding process of HIV-1 and EIAV by binding to HIV-1-p6 or EIAV-p9 of Gag (Strack et al., 2003). Clathrin forms the coat of vesicles transporting cellular cargo. Its cellular partners include adaptor proteins AP-1, -2, -3 and -4 linking clathrin to membrane lipids (Edeling et al., 2006). AP-1 and AP-2 have been shown to bind HIV-1 Env gp41 C-terminus regulating the subcellular location of Env (Berlioz-Torrent et al., 1999). AP-1μ, a subunit of AP-1, also binds Gag and silencing of AP-1μ by RNA interference showed a reduced Gag release from transfected cells (Camus et al., 2007). Since these proteins were also found in VSV-G-only samples it is not possible to confidently exclude the possibility that these proteins may also associate with budding VSV-G for example in the form of vesicles and co-purify with vectors. However, results of the publications described above, show that these proteins may play an active role in viral particle formation.

Comparison to other proteomics studies of viral vectors or enveloped virions. Of the proteins identified in all five HIV-1 Gag containing LV samples Beta-Actin, HSPA1A, HSC71, CypA, EEF1A, ENO1, and ALIX have been reported in one or more studies to be associated with wild type

HIV-1 virions by MS analysis (Chertova et al., 2006, Saphire et al., 2006, Ott, 2008) or lentiviral vectors (Wheeler et al., 2007, Denard et al., 2009). Our data is put into context of published studies on the cellular proteins previously identified by MS analysis in viral vector samples and enveloped viruses other than HIV-1. Some of the proteins that were identified in our vector samples by LC-MS/MS have also been shown to be associated with budding virions by MS/MS as listed in Table 19, supporting the possibility that some of these detected proteins play a role in vector assembly.

Table 19: LC-MS/MS Identified Proteins Common to Enveloped Viruses and Vectors

			MDM-HIV-1	VSV-New Jersey strain	Vaccinia Virus IMV	KSHV/HHV-8	HCMV	EBV	Influenza Virus	HIV-1 vector	MMLV vector
full name	UniProt #	gene name	Chertova (2006)	Moerdyk-Schauwecker (2009)	Chung (2006)	Zhu (2005)	Varnum (2004)	Johannsen (2004)	Shaw (2008)	Denard Wheeler (2009)	Segura (2008)
Actin, beta	Q53G76	ACTB									
Actin, alpha skeletal muscle	Q5T8M8	ACTA1								A1	
Tubulin, alpha 1b	Q8WU19	TUBA1B	1		1C			A3			
Tubulin, alpha 1a	Q71U36	TUBA1A									
Tubulin, beta class I	Q5JP53	TUBB									
Clathrin heavy chain 1	Q00610	CLTC									
Programmed cell death 6-interacting protein (ALIX)	Q8WUM4	PDCD6IP									
Elongation factor 1-alpha	Q504Z0	EEF1A1		1,2		2					
Alpha-enolase	P06733	ENO1									
Cyclophilin A	A8K486	PPIA									
Glyceraldehyde-3-phosphate dehydrogenase	E7EUT5	GAPDH									
Heat shock cognate 71 kDa protein	P11142	HSPA8									
Heat shock 70 kDa protein 1A/1B	P08107	HSPA1A									
Heat shock protein HSP 90-beta	P08238	HSP90AB1									
Neuroblast differentiation-associated protein (AHNAK)	Q09666	AHNAK									
Annexin A2	P07355	ANXA2	A1/2		A1/2	A2/6			A1/2		
Annexin A5	E9PHT9	ANXA5	P62937					6			
Cofilin 1 , non-muscle	G3V1A4	CFL1									

			MDM – HIV-1	VSV-New Jersey strain	Vaccinia	KSHV/ HHV-8	HCMV	EBV	Influenza Virus	HIV-1 vector	MMLV vector
Radixin	P35241	RDX									
Moesin	P26038	MSN									
Ezrin	E7EQR4	EZR									
Ras-related protein Rab-1A	B7Z8M7	RAB1A									
Ras-related protein R-Ras2	B7Z6C4	RRAS2									
Ras-related protein Rab-7a	P51149	RAB7A									
Polyubiquitin-C	P0CG48	UBC	Ubi	Ubi	Ubi					Ubi4	Ubi
Syntenin-1	B4DHN5	SDCBP									
Chloride intracellular channel protein 1	O00299	CLIC1									
Guanine nucleotide-binding protein G(k) subunit alpha	P08754	GNAI3		Go				2			
Triosephosphate isomerase	B7Z5D8	TPI1									
Histone cluster 1, H1c	P16403	HIST1H1C									
H2A histone family, member X	P16104	H2AFX									
Histone H2B	A8K9J7	HIST2H2BE									
Histone cluster 1, H2bj	P06899	HIST1H2BJ									
Histone cluster 1, H4h	Q0VAS5	HIST1H4H									
Y box binding protein 1	P67809	YBX1									
F-actin-capping protein subunit alpha-1	P52907	CAPZA1									
Plastin 3	B7Z6M1	PLS3									
Profilin-1	P07737	PFN1									
Filamin A	Q5HY54	FLNA									
Talin 1	Q5TCU6	TLN1									
Transferrin receptor protein 1	P02786	TFRC									
Myosin-9	P35579	MYH9									

			MDM – HIV-1	VSV-New Jersey strain	Vaccinia	KSHV/ HHV-8	HCMV	EBV	Influenza Virus	HIV-1 vector	MMLV vector
Prostaglandin F2 receptor inhibitor	Q4QQP8	PTGFRN									
EH domain-containing protein 4	Q9H223	EHD4		1				1			
ATPase, Na+/K+ transporting, alpha 1 polypeptide	B7Z3U6	ATP1A1									
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	E7EX29	YWHAZ									
Solute carrier family 3 (amino acid transporter heavy chain), member 2	F5GZS6	SLC3A2									
60S acidic ribosomal protein P2	P05387	RPLP2									
60S ribosomal protein L7a	P62424	RPL7A									
60S ribosomal protein L3	P39023	RPL3									
Ribosomal protein L4	P36578	RPL4									
Ribosomal protein S8/S3	Q5JR95	RPS8/S3									
Nucleophosmin	P06748	NPM1									
MARCKS-related protein	P49006	MARCKSL1									
Poly [ADP-ribose] polymerase 1	P09874	PARP1									
Basigin	P35613	BSG									
Creatine kinase B-type	P12277	CKB									
Lactadherin	Q08431	MFGE8									
Tubulin, beta 3 class III	Q9BV28	TUBB3									
Alpha-actinin-4	A4K467	ACTN4					1				
Fructose-bisphosphate aldolase A	P04075	ALDOA									

Comparison of LC-MS/MS identified host cell proteins in our vectors with published studies using MS analysis of enveloped secreted virions of: Human Immunodeficiency Virus-1 (HIV-1)

(Chertova et al., 2006); Kaposi's Sarcoma-Associated Herpesvirus (KSHV) (Zhu et al., 2005); Human Cytomegalovirus (HCMV) (Varnum et al., 2004); Epstein-Barr-virus (EBV) (Johannsen et al., 2004); Influenza Virus (Shaw et al., 2008); VSV-New Jersey strain (Moerdyk-Schauwecker et al., 2009); except vaccinia virus, which are intracellular mature virions (Chung et al., 2006); comparison with viral vectors derived from MMLV (Segura et al., 2008) or HIV-1 (Wheeler et al., 2007, Denard et al., 2009). Annotations in the green boxes indicate if another family member of an identified protein in our study was found in the referenced study.

In addition to the proteins described in the text above some of the other host proteins detected here have been identified by another group using reverse phase LC-MS/MS in sucrose-gradient in purified, CD45 depleted monocyte derived macrophages (MDM)-produced HIV-1 virions from full-length infectious molecular clone NLAD8 (Chertova et al., 2006). These host proteins include actin binding proteins Ezrin, Radixin, Filamin and Talin 1 and other transmembrane proteins (Annexin A2 and A5, RRAS2). Other proteins detected in our samples such as GAPDH and Ubiquitin were also found in wild type HIV-1 virions, purified by sucrose gradient ultracentrifugation after harvest from 293T cells or Jurkat T cells (Saphire et al., 2006a).

Interestingly only one of all LC-MS/MS detected proteins, AHNAK (or neuroblast differentiation-associated protein or desmoyokin) was unique to RD pseudotyped vectors RDpro-GFP and RDpro-Empty produced by STAR cells. AHNAK has been previously detected in MDM-produced HIV-1 virions (Chertova et al., 2006). Virions produced in Chertova's study were prepared by transfecting a HIV-1 provirus-encoding DNA plasmid into MDMs. Envelope proteins are not the same in Chertova's and our study (HIV gp160 and VSV-G Env, respectively), suggesting that AHNAK is not necessarily associated with the RDpro envelope in our samples. To date a function for AHNAK in the context of virus formation has not been described. It will be further analysed and its potential role in vector formation discussed in the following chapter (chapter 5).

Proteins only detected in VSV-G pseudotypes included GAPDH, RAB1A, Tubulin beta class 1, EH domain containing protein 4. These proteins have also been detected in a study of Vesicular Stomatitis virus (VSV) propagated in A549 human lung carcinoma cells and purified by sucrose-gradient (Moerdyk-Schauwecker et al., 2009) indicating that these cellular proteins might associate with VSV-G Env specifically.

In the study by Chung et al. (2006) vaccinia virus was propagated in HeLa cells, intracellular mature virions (IMV) were extracted from homogenised cells by sucrose- gradient ultracentrifugation and twenty-three host proteins were detected using MS, of which 12 are also detected in our study. Sixteen out of 21 proteins identified by LC-MS/MS in virions of Kaposi's Sarcoma-Associated Herpesvirus (KSHV), also known as human herpes virus 8 (Zhu, et al 2005), are shared with our study. KSHV virions were grown in the cell line BCBL-1, a primary effusion lymphoma cell line latently infected with KSHV, and purified by double gradient ultracentrifugation. In virions of Human Cytomegalovirus (HCMV) a total of 71 host proteins were identified by MS/MS of which 24 were also found in our samples (Varnum, 2004). Virions were grown in human dermal fibroblasts, followed by purification using sorbitol cushion ultracentrifugation and nycodenz gradient, a nonionic iodinated gradient medium. Forty-eight host proteins were identified by LC-MS/MS in virions of the gammaherpesvirus Epstein-Barr-virus (EBV) after purification by dextran-ultracentrifugation and deglycosylation. Twenty-one out of 48 identified proteins were also detected in our samples, including the most frequently detected proteins tubulin, actin-binding proteins and cofilin (Johannsen, 2004). MS analysed Influenza virions, propagated in MDCK cells canine kidney cells, purified by sucrose gradient ultracentrifugation and deglycosylated, shared 9 proteins out of 16 detected cellular proteins with our vectors, such as Beta-Actin, Tubulin alpha and beta, Cofilin, Profilin-1 (Shaw, 2008). To date there is no published study of RD114 virion protein composition.

Mass spectrometry has also been used for host protein identification in retroviral vector samples. Moloney murine leukaemia virus (MMLV) vector particles pseudotyped with VSV-G

were harvested from the 293-GPG packaging cell line, concentrated by diafiltration and purified by rate zonal ultracentrifugation followed by an extra step, purification by SEC (Segura et al., 2008). Out of 25 identified host proteins 15 proteins were identified that were also found in our samples. Most of them have been identified in at least one other virus shown in Table 19.

The overlap in identified cellular proteins between our and some of the studies on enveloped viruses or vectors is up to 50% and covers mainly proteins of the cytoskeleton or cytoskeleton regulators but also proteins associated with vesicular trafficking and transmembrane proteins. Functional categories of cellular proteins found in SEC purified vector samples are similar to the ones identified in virions purified by sucrose gradient. This could be due to similarities in assembly and budding processes in different enveloped viruses that lead to association with similar cellular proteins, for example secretory pathway trafficking of envelope proteins.

Alternative methods have been described to find links of host cell proteins to HIV-1 replication. Screens using siRNA to knock down cellular expression of human genes in HIV-1 infected HeLa or 283T cells were used to analyse differences in infectious particle production and particle infectivity (Brass et al., 2008, König et al., 2008, Zhou et al., 2008). Out of 842 identified genes in all three screens, 34 were identified in at least two screens. There is no overlap of these 34 proteins with proteins detected by MS in our study. Variations in the experimental set up of the three screens are thought to cause the small overlap of candidate genes common to all three siRNA screens. Similarly the difference in procedures to our project does not allow simple comparison of data.

Less cellular proteins in RDpro compared to VSV-G pseudotypes. Mass spectrometry of vector samples showed less cellular protein species in RDpro-pseudotyped vectors compared to transiently produced vectors when analysing a comparable number of viral particles. A third set of stably produced samples was analysed (MS set 3) containing similar levels of p24 viral protein, 10 µg of total protein, to make RDpro-pseudotyped samples comparable to transiently

produced MS samples regarding their particle numbers. Results showed that samples produced by STAR cells contain less cellular proteins than transiently produced samples even when a comparable number of viral particles were analysed. To confirm that RDpro-pseudotyped vectors truly contain less host proteins, a transiently produced sample of RDpro-pseudotyped vectors would have had to be included in this study. Similar to RDpro-pseudotypes, Gag/Pol-GFP vector samples, without a viral envelope, contained substantially fewer host cell proteins than VSV-G pseudotypes. This could be due to the possibility that more cell derived vesicles are present in VSV-G pseudotyped vectors and were co-purified with vector particles in transiently produced samples.

VSV-G vesicles. VSV-G vesicles have been shown to co-purify with vector preparations by continuous sucrose gradient ultracentrifugation produced by 293T cells (Pichlmair et al., 2007). Cell membrane derived vesicles have been identified in wild type HIV-1 virion preparations from T cells that co-purified along HIV-1 virions by sucrose- gradient ultracentrifugation. These vesicles were found to be of similar density to HIV-1 virions and between 50 to 500 nm in size compared to 100 nm HIV-1 virions (Gluschankof et al., 1997, Bess et al., 1997). It has been shown that treatment with the serine protease subtilisin digests proteins on the surface but not inside the virions (Ott et al., 1995) and in combination with subsequent sucrose-gradient centrifugation reduced the presence of proteins in the vector particle sample that are potentially derived from vesicles (Ott et al., 1996). Discrimination between host cell proteins incorporated in virions and externally bound or co-purified proteins can also be achieved by using another serine protease, proteinase K, treating vectors before MS analysis (Denard et al., 2009, Moerdyk-Schauwecker et al., 2009). Denard et al. (2009) identified the following proteins after protease K treatment, hence they are thought to be inside virions, which are shared with proteins found in our study: Beta-Actin, Clathrin, HSPA1A, HSC71, Synthenin, GAPDH, CypA, EEF1A, ENO1, Annexin A2 and A5, Ezrin as well as ALIX. However, complete elimination of co-purified vesicles was not achieved when using subtilisin digests (Ott et al., 1995) hence this method may not be used for confident location of proteins in vector samples.

Thus, at least some of the 38 proteins only detected in VSV-G pseudotyped sample, described in the results of this study, could originate from host proteins associated with VSV-G or potentially with co-purified VSV-G vesicles. Protease digestion in combination with a method that visualises candidate surface proteins such as immunogold labelling could verify if a protein is associated with the vector surface or inside the vectors (Segura et al., 2008).

***Specificity of Protein Association*.** Based on our results and those of other vector and virus studies, different viral particles seem to have commonly associated cellular proteins regardless of the production system or the vector pseudotypes. Gurer et al. (2005) showed that co-transfection of producer cells with DNA plasmids expressing HIV-1 provirus and a mutant form of HSP70 completely abolished HIV-1 virion formation. Hence the association of some of the here identified host cell proteins that were also identified in above described studies of MS analysed wild type viruses might be due to an essential role in assembly and budding of vectors or viruses. On the other hand of all the here identified proteins in particular those highly abundant in the cell, especially at the plasma membrane, the vector assembly site (for example actin and tubulin), could be incorporated into viral particles purely or partly due to their presence. A study analysed HIV-1 cores, HIV-1 virions without envelope proteins, by LC-MS/MS. These HIV-1 cores were purified from different cell types, including an infected T cell line and THP1 cells (a model of monocytes) as well as activated THP1 cells (a model of MDMs) by sucrose-gradient ultracentrifugation and a 1% Triton X-100 layer (Santos et al., 2012). Classification of the 202 identified cellular proteins showed that the majority are RNA-binding proteins, components of the cytoskeleton and cytoskeleton regulators, chaperones, DNA-binding proteins, proteins involved in vesicular transport and components of the ubiquitin-proteasome system, similar to the major categories of the proteins identified in our study. The authors show that proteins of the cytoskeleton are incorporated in the different HIV-cores and hypothesise that this is due to their high abundance in the cell and their location relative to the assembly site increasing the likelihood of their acquisition by virions during assembly and budding (Santos et al., 2012). However, high abundance at viral assembly site and an actual

role in assembly is not mutually exclusive as above discussed publications show where roles of actin and tubulin in HIV-1 assembly (Jolly et al. 2007) and specific binding of HIV-1 Gag to actin (Wilk et al., 1999) have been documented. Also, overlap of results of our study with a study on HIV-1 virions produced in a different cell type, MDMs, compared to 293T cells used here, suggests that at least a subset of these host cell proteins is not randomly incorporated (Chertova et al., 2006).

Technical aspects. Surprisingly, in any of the RDpro-pseudotypes, RDpro Env was only detected when using the rejection method in LC-MS/MS, the method that selects also less abundant peptides for MS/MS analysis. This resulted in detection of one peptide of RDpro Env protein confidently and up to three peptides at low confidence covering 4 to 7 % of the whole RDpro Env protein sequence. It has been suggested that the wild type envelope of RD114 does not incorporate efficiently into lentiviral vectors and it was shown that a modified form of the envelope can incorporate more stably on to vector particles (Sandrin et al., 2002). Strang et al. (2003) documented that the infectious titer of RDpro-pseudotyped vectors after concentration by ultracentrifugation could be increased by reducing centrifugation speed whereas lower spin speeds did not increase VSV-G pseudotyped vector titer, indicating that RDpro Env is more fragile than VSV-G Env. Due to lower infectious titer of LVs pseudotyped with wild type RD114 Env, modified versions of the protein have been designed before. RDpro, used in this study, was created by replacing the R peptide cleavage site with the HIV-1 matrix/capsid (MA/CA) protease cleavage sequence (Ikeda et al., 2003). This increased the infectious titer of LVs pseudotyped with RDpro compared with RD114 but it is not known if it also increased the stability, hence the here used RDpro might not be efficiently incorporated on to vector particles resulting in reduced presence of RDpro peptides in MS vector samples.

Alternatively, the lack of detected RDpro Env peptides could be due to the glycosylation of RDpro envelope protein. MS analysis used here did not allow the detection of peptide glycosylation as MS study of protein or peptide glycosylation requires specialised instrument

settings. Hence this modification was not included in the protein database. Generally, potential reasons for not detecting glycosylated peptides are the lower efficiency of ionisation of glycosylated compared to non-glycosylated peptides (Ruiz-May et al., 2012). Alternatively, glycosylated peptides exceed the mass scan region. High-molecular weight modifications, caused by glycosylation, can shift the molecular weights of the glycosylated peptides out of the mass scan region in which they can be sequenced (Parker et al., 2010), hence glycosylated peptides can be difficult to detect by MS/MS. VSV-G Env and RDpro potentially have many N- and O-glycosylation sites, as predicted by Expasy-GlcoMod (http://web.expasy.org/glycomod/), hence many RDpro Env peptides could be heavily glycosylated. The calculated molecular weight, using Expasy-ProtParam (http://web.expasy.org/protparam), based on the amino acid sequence of RD114 SU (amino acid position 18 to 509) (Sandrin et al., 2002) is 53920.7 Da. However, the migration pattern of the RD114 SU in polyacrylamide gels indicates an actual size of 7000 Da (Strang et al., 2004), hence the additional mass of 16079.3 Da is likely to be derived from glycans attached to RD114 SU. A way to potentially improve RDpro Env detection is to treat purified vector samples with N- and O-glycanases, so that more deglycosylated peptides from RDpro Env are available for sequencing using our experimental conditions. Johannsen et al. (2004) reported detection of Epstein-Barr virus (EBV) glycoprotein only after EBV samples were deglycosylated using N- and O-glycanases. However, deglycosylation does not always result in the detection of increased protein species. A comparison of MS identified host protein of glycosylated and deglycosylated influenza virus showed that the same number of protein species were detected with and without sample deglycosylation (Shaw et al. 2008). None of the 17 different VSV-G peptides we detected in VSV-G pseudotyped vector samples have an N-glycosylation motif, as predicted by Expasy-GlcoMod (http://web.expasy.org/glycomod/). However they are all located between positions 175 to 488 of the amino acid sequence, part of the VSV-G ectodomain and contain serine or threonine residues, potential sites for O-glycosylation, hence the presence of

potential glycosylation sites does not necessarily mean these sites are heavily glycosylated and prevent peptides from being detected by MS.

MS analysis of our samples did not detect the myristoylated N-terminal amino acid glycine of MA of HIV-1 Gag (Veronese et al., 1988), supporting the assumption that in our samples post-translational modifications hinder MS detection of peptides. The peptide mass that contains this myristoylated amino acid might be out of the predicted mass region as myristate adds a molecular mass of 228 Da (NCBI, pubchem, http://pubchem.ncbi.nlm.nih.gov/).

Aspects to consider in future work. For complete MS sample analysis all vector samples should be analysed using the same conditions, specifically using a comparable number of vector particles (as was used in all viral vector samples of MS set 3), the same data acquisition method, such as the rejection method and the same data analysis method (minimum amount of peptides required per protein). This would potentially detect low abundance peptides in VSV-G pseudotypes that were not detected using the normal acquisition method, allowing a more accurate estimation of how many more proteins can be detected in VSV-G-pseudotypes compared to RDpro-pseudotypes.

Data of identified host proteins must be treated with caution and association of host proteins with vectors should be confirmed by other methods such as knock-down of protein expression in producer cell followed by titration of produced vectors, however most of the proteins described above are essential to the function of the producer cell, for example proteins of the cytoskeleton or other proteins encoded by housekeeping genes. Effects of a complete knock-down of these host proteins on vector production may not exclusively be accounted to the missing interaction of host and viral proteins but could be also caused by the reduced housekeeping function of the cellular protein. An interaction of host and viral protein may also be determined using a co-immunoprecipitation assay in cell lysates to isolate a specific host protein and analyse if it is bound to a viral protein. Furthermore an additional sample could be analysed such as the supernatant of untransfected 293T cells that is concentrated in the same

way as vector samples in order to potentially identify which of the MS detected proteins in our vector samples are unrelated to vector production.

Future work will focus on proteins that are common to all vector samples and are less well characterised and for those reasons may be considered relevant in vector particle assembly and budding from the producer cell.

5. Effects of Cellular Protein Expression on Producer Cells in Lentiviral Vector Production

5.1. Introduction

The retroviral genome encodes only a limited number of proteins making an extensive interaction of replicating viruses with the host-cell machinery essential. Host proteins interact with viral proteins throughout the HIV-1 replication cycle. Especially interesting for this study are interactions of cellular and viral proteins during the late stages of HIV-1 virus replication that are related to vector production including viral protein expression, transport to the assembly site and budding. Among the reported host-virus interactions is the recruitment of the endosomal sorting complex (or 'endosomal sorting complex required for transport', ESCRT) by p6 of HIV-1 Gag to support budding of immature viral particle from the infected cell (Balasubramaniam and Freed, 2011). It also has been shown that matrix (MA), part of Gag binds to the cell membrane as an important step in virus assembly (Saad et al., 2006). HIV-1 envelope protein precursor gp160 is synthesized and glycosylated in the rough endoplasmic reticulum (RER) before cleavage to the mature glycoprotein gp120 by cellular proteases such as furin (Checkley et al., 2011, Hallenberger et al., 1992). Lentiviral vectors (LVs) used in this study are derived from HIV-1 virus hence may depend on host cell protein interaction in a similar way to wild type HIV-1. The mass spectrometry (MS) detection of host cell proteins

associated with purified virus particles and viral vectors is one approach to identify host factors that interact with the virus or vector and play a role in their generation within the cell. Proteins that were found in purified vector samples in this project have also been reported in one or more studies to be associated with wild type HIV-1 virions, such as the transcription factor 1-alpha (EEF1A1), programmed cell death 6-interacting protein (ALIX or AIP1), annexin A2 or 5 and alpha-enolase (Chertova et al., 2006, Saphire et al., 2006, Ott, 2008) and in studies on crude or purified LVs (Wheeler et al., 2007, Denard et al., 2009).

The main aim of the project presented here was to improve the understanding of fundamental virus-host interaction in viral vector assembly, in particular, the protein-protein and protein-RNA interactions during viral particle assembly and formation. Mass spectrometry of purified vector samples identified certain proteins that are particularly interesting for further analysis of their function or potential role in LV assembly. This chapter focuses on proteins that are less well characterised regarding a role in HIV-1 replication and were detected in most of the analysed vector samples or unique to a specific vector sample, hence may be relevant in vector particle assembly and budding from the producer cell. The focus of this project will include the following proteins. AHNAK (alternatively neuroblast differentiation-associated protein or desmoyokin) is unique to STAR cell produced vectors in this study. AHNAK has been described the first time in desmosomes of bovine muzzle epidermis (Shtivelman et al., 1992, Hieda and Tsukita, 1989). It was found to be distributed in the cytoplasm and nuclei in keratinocytes (Hashimoto et al., 1995). It locates to the plasma membrane when bound to protein kinase C (PKC) (Hashimoto et al., 1995, Lee et al., 2008) and is thought to activate protein kinase C-alpha (Lee et al., 2008) suggesting its implication in cell signalling. Other studies identified it as an activator of phospholipase C-γ (Sekiya et al., 1999). In cardiomyocytes AHNAK was shown to connect actin in the cytoskeleton to the ß2a-subunit of cardiac L-type Ca^{2+} channels (Hohaus et al., 2002). It has also been found in cytoplasmic vesicles which undergo exocytotic fusion after calcium increase (Borgonovo et al., 2002). To date AHNAK has not been associated with a role

in viral replication but it has been detected as a host cell protein in monocyte derived macrophage- (MDM) produced HIV-1 virions (Ott et al., 1996, Chertova et al., 2006) and was recently highlighted in a study screening for HIV-1 responsive phosphoproteins that impact on HIV-1 replication in MAGI cells (Wojcechowskyj et al., 2013). MAGI cells are also called HIV indicator cells and derived from HeLa cells expressing CD4, CCR5 and ß-galactosidase under the control of a truncated HIV-1 LTR. HIV-1 can be only expressed in the presence of Tat, thus expression levels can be directly related to HIV-1 expression levels (Kimpton and Emerman, 1992). AHNAK knock-down mediated by small interfering RNAs (siRNA) resulted in reduced HIV-1 entry in MAGI cells between three to 25 fold compared to the 'non-targeting siRNA' control (Wojcechowskyj et al., 2013). One of the proteins identified by mass spectrometry (MS) in all purified vector samples is elongation factor 1-alpha (EEF1A1). It is known for binding and transport of aminoacyl-tRNA to the ribosome as part of translation of RNA into amino acid sequence. It also associates with HIV-1 reverse transcriptase and integrase in reverse transcription (Warren et al., 2012) and has been detected in HIV-1 virions (Saphire et al., 2006a) and HIV-1 derived LVs (Wheeler et al., 2007, Denard et al., 2009). Another protein detected in all vector samples was alpha-enolase (ENO1). It is of interest for further analysis as it has been shown to interact with the Marek's Disease Virus (MDV), a herpesvirus causing a lymphoproliferative disease of chickens (Niikura et al., 2004) and was also detected in MDM-produced HIV-1 virions (Chertova et al., 2006) as well as LVs (Denard et al., 2009). Myristoylated alanine-rich C-kinase substrate-related protein 1 (MARCKSL1) is one of the two members of the MARCKS protein family and is involved in membrane-cytoskeletal signalling, such as cell migration, adhesion and secretion and phagocytosis. It was identified in all vector samples apart from non-enveloped Gag/Pol-GFP vectors. The second family member, MARCKS, was identified in VSV-G pseudotyped samples only and has been reported to be 30 fold upregulated in Eppstein-Barr virus infected cells of Burkitt's lymphoma (BL) (Birkenbach et al., 1993). It has not yet been associated with HIV-1 however it is still an interesting candidate for further studies. One of the proteins detected in MS analysis and common to all purified vector

samples was ALIX (ALG-2-interacting protein X or programmed cell death 6-interacting protein PDCDI6P). It has been connected to endosome trafficking and plays a role in multivesicular bodies biogenesis, binding to charged multivesicular body protein 4 (CHMP4) (Katoh et al., 2003). ALIX has also been detected in proteome studies of HIV-1 virions, LVs and gamma-retroviruses (RVs) (Chertova et al., 2006, Wheeler et al., 2007, Denard et al., 2009, Segura et al., 2008). Instead Tumour susceptibility gene (TSG101) is thought to be important for in HIV-1 budding. It is not one of the proteins that was identified in MS analysis of LVs in this project but it is another component that is involved in viral budding through the ESCRT machinery and is a known factor of the ESCRT-I complex (Babst et al., 2000).

To interrogate the function of a protein one approach is to reduce or abolish its expression and study subsequent effects on a specific phenotype. The use of RNA interference (RNAi) leads to sequence specific knock-down of protein expression by reducing RNA levels in the target cell which in turn results in reduced protein expression. To study host-virus interactions this technique can be used on an individual protein basis as presented here or in cell based high-throughput screening methods and has been applied to identify host-factors for HIV-1 infection in 293T (Brass et al., 2008, König et al., 2008) and HeLa cells (Zhou et al., 2008, Brass et al., 2008) as well as in the T-cell line, Jurkat (Kok et al., 2009). None of the candidate proteins AHNAK, EEF1A, ENO1, MARCKSL1 or ALIX had been identified as host cell factors in these studies, however the authors do not provide full lists of all transcripts that were screened. Neither are genes listed whose transcript expression did not influence HIV-1 infectivity. Zhou et al. (2008) used TSG101 as a positive control. They depleted by siRNA TSG101 expression in 293T cells, infected with proviral HIV-1 and transduced indicator cells with the produced virus which showed significantly reduced the infectivity. In this chapter it was attempted to knock-down protein expression of MS identified proteins AHNAK, EEF1A1, ENO1, MARCKSL1 and ALIX as well as TSG101 in producer cells of both transient and stable systems by delivering small hairpin RNA (shRNA) using LVs to the packaging cells to study

effects on LV production and infectivity. The principle of shRNA is based on the RNAi or RNA silencing, whereby double-stranded RNA (dsRNA) induces the homology-dependent degradation of cognate mRNA. This principle was first discovered in nematodes (Fire et al., 1998) followed by the discovery in plants (Hamilton and Baulcombe, 1999) and is well preserved across different species.

There are at least three classes of endogenous RNAs: micro RNAs (miRNAs), endogenous small interfering RNAs (endo-siRNAs) and PIWI interacting RNAs (piRNAs). PiRNAs or repeated-associated small interfering RNAs (rasiRNAs) were found first in *D. melanogaster* expressed in germ-line cells regulating the silencing of transposons (Aravin et al., 2003). Micro RNA and endo-siRNA are mainly post-transcriptional regulators by base-pairing with their specific target mRNA leading to translational repression and exonucleolytic mRNA decay. The first miRNA, lin-4, has been identified in *C.elegans* (Lee et al., 1993). A miRNA primary precursor RNA (pri-RNA) is transcribed in the nucleus forming a stem loop which is recognised by Drosha and cropped to form precursor RNA (pre-RNA) (Lee et al., 2002, Lee et al., 2003). The hairpin is then exported to the cytoplasm by exportin 5 (EXP5) (Yi et al., 2003). A dsRNA-specific endonuclease (DICER) digests pre-RNA into 22 nucleotide long dsRNA molecules (Bernstein et al., 2001) that are loaded onto a nuclease complex to form an RNA-induced silencing complex (RISC) (Hammond et al., 2000). The guide strand for the dsRNA remains in the complex whereas the passenger strand gets degraded. The RISC then targets the complementary endogenous gene transcript by base-pairing and cleaves the mRNA, leading to reduced protein translation and in return decreased protein expression.

For *in vitro* knock-down of protein expression a shRNA can be engineered. ShRNAs are based on endogenous miRNAs. A single strand RNA folds into a hairpin shape and is then cut into a shorter dsRNA (Paddison et al., 2002). To deliver them to the target cell they are cloned into a lentiviral vector genome plasmid that is then packaged into LV particles (shRNA-LVs). Upon transduction of the target cells with shRNA-LVs the shRNA sequence is integrated into the cell

genome and stably expressed. Vector genome plasmids used in these experiments are GIPZ lentiviral shRNAs purchased from Thermo Scientific. The plasmid contains the LV genome elements 5' and 3' LTR, central polypurine tract (cPPT) to increase transcription, the packaging sequence psi (Ψ) and the Woodchuck hepatitis virus posttranscriptional regulatory element (WPRE or WRE) to enhance expression of the transgene, in this case of the shRNA. The shRNA is based on the endogenous miRNA 30 (mir-30) providing efficient processing in the target cell. To mark shRNA expression the reporter gene turbo green fluorescent protein (tGFP) driven by the cytomegalovirus (CMV) promoter is included and placed upstream of an internal ribosome entry site (IRES) that drives the expression of an antibiotic selection marker, puromycin (Figure 20) and the shRNA.

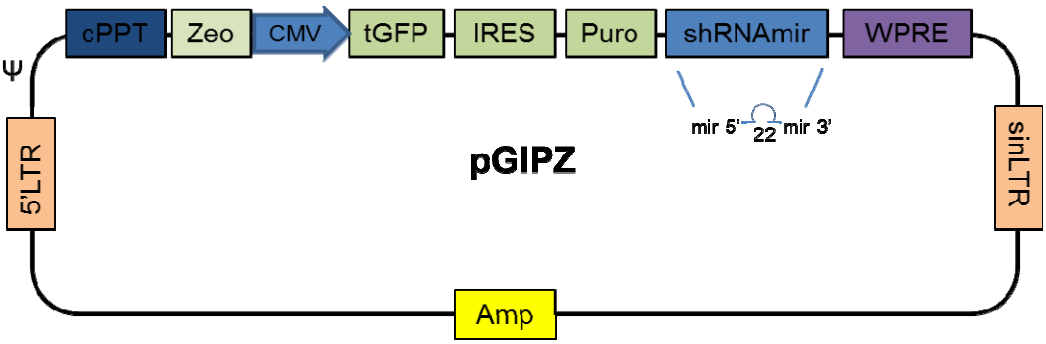


Figure 22: Human pGIPZ shRNAmir lentiviral genome plasmid (Thermo Scientific).Vector elements: long terminal repeats 5' and 3' LTR, central polypurine tract (cPPT), packaging sequence (psi,Ψ), Woodchuck hepatitis virus posttranscriptional regulatory element (WPRE or WRE), short hairpin RNA based on miRNA-30 (shRNAmir), turbo green fluorescent protein (tGFP) driven by cytomegalovirus (CMV), internal ribosome entry site (IRES), puromycin (puro).

5.2. Aim

The aim of this chapter was the analysis of selected host proteins that were identified in purified lentiviral vector samples by MS, for their potential effect on vector production from producer cells. SiRNA was applied to knock-down expression of one selected protein at a time in packaging cells. Knock-down levels of several candidate proteins were screened and proteins with considerable reduction of expression levels selected for further study. Vectors produced by knock-down-producer cells were characterised by quantification of numbers of infectious particles as well as physical particle numbers before and after protein knock-down.

5.3. Results

Lentiviral vectors (GIPZ-LVs) carrying a shRNA sequence were prepared by co-transfection of 293T cells with DNA plasmids expressing HIV-1 *gag-pol* (p8.91) and VSV-G envelope (pMDG) as well as a pGIPZ plasmid carrying a shRNA targeting AHNAK, EEF1A1, ENO1, MARCKSL1, ALIX or TSG101 (Thermo Scientific). The negative shRNA control GIPZ scramble was used consisting of a shRNA sequence of random nucleotides that does not code for any known mammalian genes. Repeat-experiments are referred to as 'sets' of samples. For analysis of infectious titers vectors of each knock-down producer and control vectors were harvested from triplicate wells. Physical titers and protein expression were analysed from single or duplicate wells. Statistical analysis was carried out for assays in which samples were analysed in duplicate or triplicate per set.

5.3.1. Screen for Efficient Knock-Down of shRNAs Targeting MS Identified Host Proteins

Several shRNAs were tested targeting different sequence parts of a host cell protein. 293T cells were transduced with a GIPZ-LV and selected in puromycin. Antibiotic selection kills the cells that do not express shRNA as GIPZ encodes a puromycin resistance gene, expressed along with the shRNA from the internal ribosome entry site (Figure 20). After at least ten days in puromycin selection medium cell lysates were analysed by western blotting for expression of the specific host protein. Knock-down of EEF1A and MARCKSL1 expression was weak even with the most effective shRNA #4 and #20, respectively (Figure 21A and B). ENO1 expression was still relatively high compared to the control after transduction with any of the GIPZ-ENO1-LVs #40, #17 or #37 (Figure 21C). Alternative methods of knock-down for these proteins should be explored. The highest reduction in AHNAK expression (Figure 21D) was seen using shRNA #6, #38 and #39. ALIX expression (Figure 21E) was mostly diminished by shRNA #18 and TSG101 protein levels were reduced best by shRNA #46 (Figure 21F). The negative knock-down control were lysates of 293T cells that were not transduced with any LV-GIPZ ('no shRNA') or LV-GIPZ-scramble. The shRNA with the highest level of knock-down, that is lowest western blot signal compared to the control for each host protein, including shRNA #6 for AHNAK, #18 for ALIX and #46 for TSG101, was selected and used in subsequent experiments. As knock-down of ALIX, AHNAK and TSG101 showed the good knock-down results compared to EEF1A, MARCKSL1 and ENO1, only three proteins were selected for further analysis.

Examination of AHNAK, ALIX and TSG101 knock-down effects on vector production is presented in this chapter. AHNAK is unique to stably produced vector pseudotyped with the RDpro envelope from the stable packaging cell line STAR, hence it was of interest to investigate if reduced AHNAK expression influences vector production in STAR cells differently to the production of VSV-G pseudotyped vectors produced in 293T cells. ALIX as one of the proteins detected by MS in all vector samples and with its known function regarding HIV-1 budding was another interesting candidate. Knock-down of TSG101 in HIV-1 infected cells had been shown

to have reduced the infectivity of produced HIV-1 virions before (Martin-Serrano et al., 2003, Zhou et al., 2008) and was expected to serve as positive control for vector production.

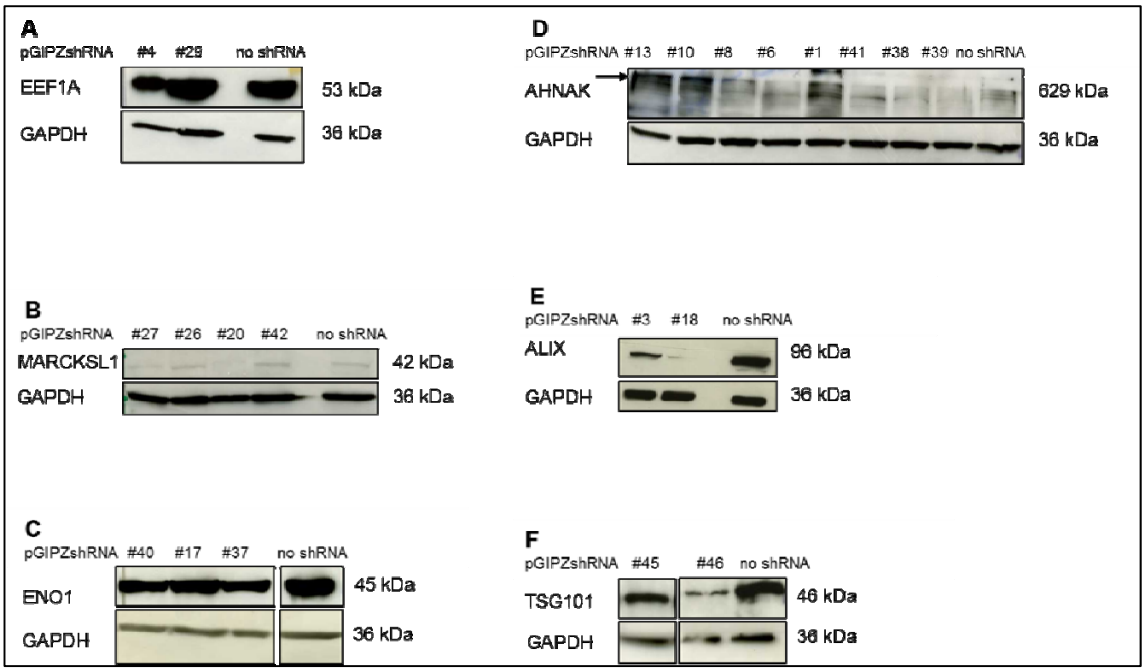


Figure 23: Knock-down efficiency of shRNA in 293T-GIPZ cell lysates. Knock-down levels of host proteins EEF1A (A), MARCKSL1 (B), ENO1 (C), AHNAK (D), ALIX (E) and TSG101 (F) were analysed. For each host protein different shRNA targets were tested to identify the one with highest knock-down efficiency (weakest signal strength compared to the control). Negative control were 293T cells without transduced GIPZ-LV ('no shRNA'). ShRNA targets AHNAK #6, ALIX #18 and TSG101 #46 were selected for subsequent knock-down experiments.

5.3.2. Analysis of GIPZ-LV producer cells

To deliver each host cell protein specific shRNA to the producer cells, STAR-RDpro-pHV (STAR) and 293T cells were transduced with GIPZ-LVs at MOI 10. Twenty-four hours after transduction, puromycin selection was started and cells remained in puromycin conditioned medium throughout the study to maintain the expression of shRNA over time. Cultures were kept at least ten days in selection medium before the start of vector production.

5.3.2.1. STAR-GIPZ LV producer cell analysis

In order to confirm knock-down of protein expression in STAR-GIPZ producer cells, the first set of STAR cell lysates was analysed at day 16 and 22 after transduction with GIPZ-LVs, representing set I. The second set of knock-down STAR cells was analysed on day 10 and 17 (representing set II). At each time point western blotting of AHNAK and TSG101 (Figure 22) or AHNAK and ALIX (Figure 23) was carried out showing significant reduction in protein expression. Negative controls were STAR-GIPZ-scramble cells and STAR cells that expressed a shRNA targeting another host protein as well as STAR cells that had not been transduced with GIPZ-LVs ('no shRNA').

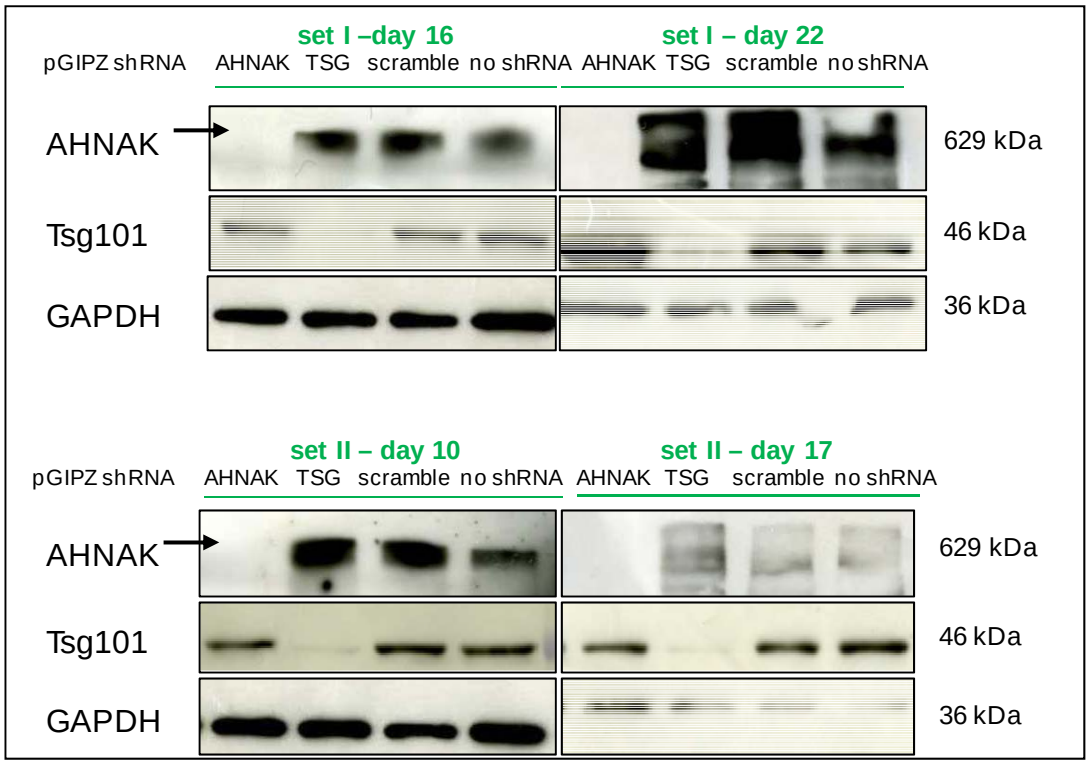


Figure 24: shRNA mediated knock-down of AHNAK and TSG101 protein expression in STAR cells. Assessment of knock-down efficiency of AHNAK and TSG101 in STAR-GIPZ cell lysates in set I and II. STAR cells were transduced with GIPZ-LVs and selected in puromycin for at least 10 days before cell lysates were analysed by western blotting of host cell proteins AHNAK, TSG101. Equal loading of total protein was verified by western blotting of GAPDH, 10 µg of total protein of cell lysates per lane were loaded. Negative control were STAR cells without transduced GIPZ-LV ('no shRNA'). Protein molecular sizes are indicated.

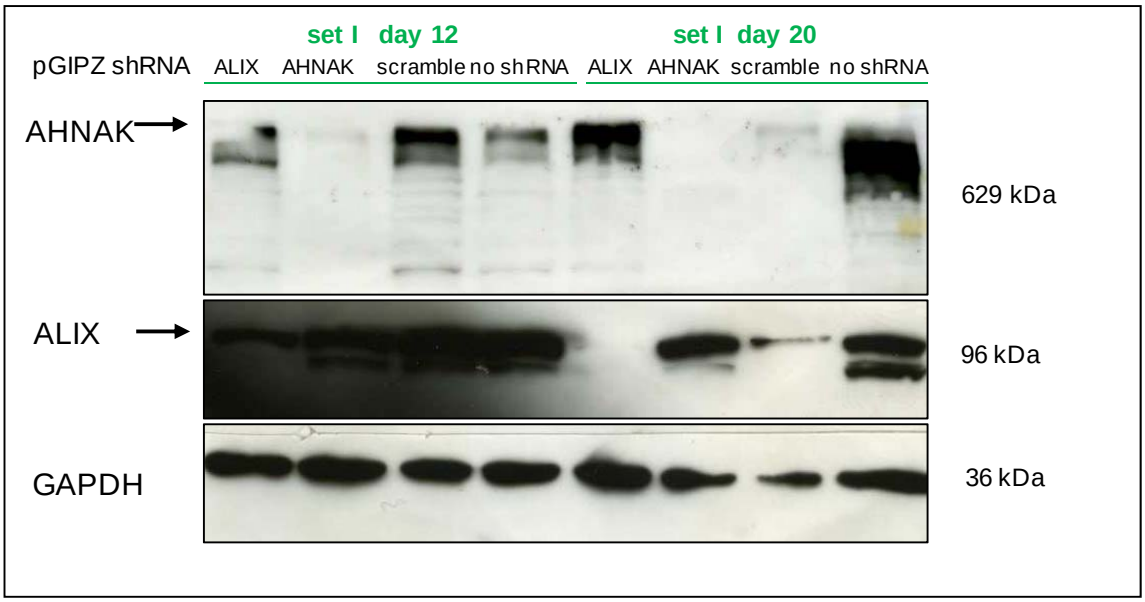
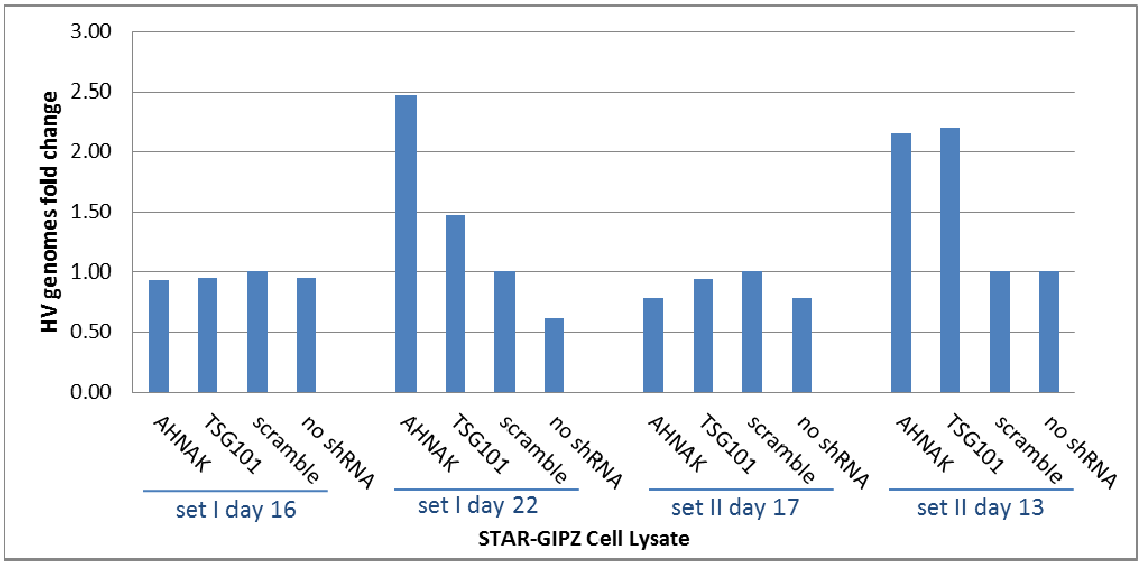


Figure 25: shRNA mediated knock-down of ALIX and AHNAK protein expression in STAR cells. Assessment of knock-down efficiency of AHNAK and ALIX in STAR-GIPZ cell lysates in set I. STAR were transduced with GIPZ-LVs and selected in puromycin for at least 10 days before cell lysates were analysed by western blotting of host cell proteins AHNAK and ALIX; equal loading of total protein was verified by western blotting of GAPDH, 10 µg of total protein of cell lysates per lane loaded. Negative control were STAR cells without transduced GIPZ-LV ('no shRNA'). Protein molecular sizes are indicated.

Following the knock-down of selected host proteins, effects on stable vector production by STAR-GIPZ cells were analysed. Effects on expression of vector RNA and proteins within the producer cells are described in this section. To determine if knock-down of candidate host cell proteins influenced vector RNA levels of the stably expressed transfer vector HV inside the cell, HV vector RNA copy fold changes in producer cell lysates from STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 compared to RNA genome levels STAR-GIPZ-scramble were measured by Q-RT-PCR, normalised to human beta-actin copies per cell (see section 2.4.4.4). To ensure specific amplification of vector genome HV, primers targeting the cPPT-SFFV promoter boundary upstream of GFP, were used. Across four repeat-experiments viral RNA copies in producer cells STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 ranged from 0.62 to 2.5 fold change compared to STAR-GIPZ-scramble in two experiments. Per experiment one well of cell per knock-down and

control sample were analysed, hence no statistical analysis was possible. RNA copy fold change in STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 producer cells compared to STAR-GIPZ-scramble was almost unchanged in set I day 16 and set II day 10. Overall no consistent difference in RNA levels of any of the samples compared to the scramble control was observed (Figure 25).



set	day	GIPZ	sample ID	Ct beta actin	Ct SFFV	SFFV fold change in host cell KD samples compared to scramble 2^-deltadeltaCt
I	16	AHNAK	179	26.0	26.0	0.94
I	16	TSG101	180	24.9	24.9	0.96
I	16	scramble	181	28.3	28.2	1.00
I	16	no shRNA	182	31.9	31.9	0.96
I	22	AHNAK	183	17.4	22.4	2.48
I	22	TSG101	184	17.8	23.6	1.48
I	22	scramble	185	17.3	23.7	1.00
I	22	no shRNA	186	17.8	24.8	0.62
II	17	AHNAK	187	27.2	27.4	0.78
II	17	TSG101	188	25.9	25.9	0.95
II	17	scramble	189	28.2	28.0	1.00
II	17	no shRNA	190	29.3	29.6	0.79
II	13	AHNAK	196	17.4	23.1	2.15
II	13	TSG101	197	17.6	23.3	2.20
II	13	scramble	198	19.7	26.5	1.00
II	13	no shRNA	199	20.8	27.7	1.00

Figure 26: No consistent effect of shRNA mediated knock-down of AHNAK and TSG101 on the vector RNA genomes shown as fold change compared to scramble control (=1) in four repeat-experiments of Q-RT-PCR in STAR-GIPZ. The viral RNA genome copy fold change in STAR-GIPZ-AHNAK, -ALIX and –TSG101 was normalised to STAR-GIPZ-ß-actin. N=3 for Q-PCR, n = 1 for sample replicates.

Western blots of viral protein p24 in producer cell lysates were analysed to assess if host cell RNA knock-down had an effect on vector production at the stage of viral RNA transcription and expression. This method allowed the semi-quantitative comparison of p24 levels of STAR-GIPZ cells to each other. Results for western blotting of one well of producer cell lysate of each day are shown in Figure 25. P24 levels in producer cells of set I showed that these are slightly increased in STAR-GIPZ-TSG101 and STAR-GIPZ-AHNAK compared to STAR-GIPZ-scramble p24 levels. This was confirmed only for STAR-GIPZ-TSG101 producers of set II. The overall differences in p24 levels could be due to minor variability between experiments. This semi-quantitative method can also detect p55 Gag protein. Levels of p55 are less varied suggesting AHNAK, ALIX and TSG101 knock-down did not affect Pr55 Gag-precursor expression. Increased levels of p24 in STAR-GIPZ-TSG101 could be due to modifications in Pr55 Gag processing after TSG101 knock-down. Taken together with the results of vector RNA copy analysis in producer cells knock-down of AHNAK and TSG101 expression has no major impact on viral protein expression.

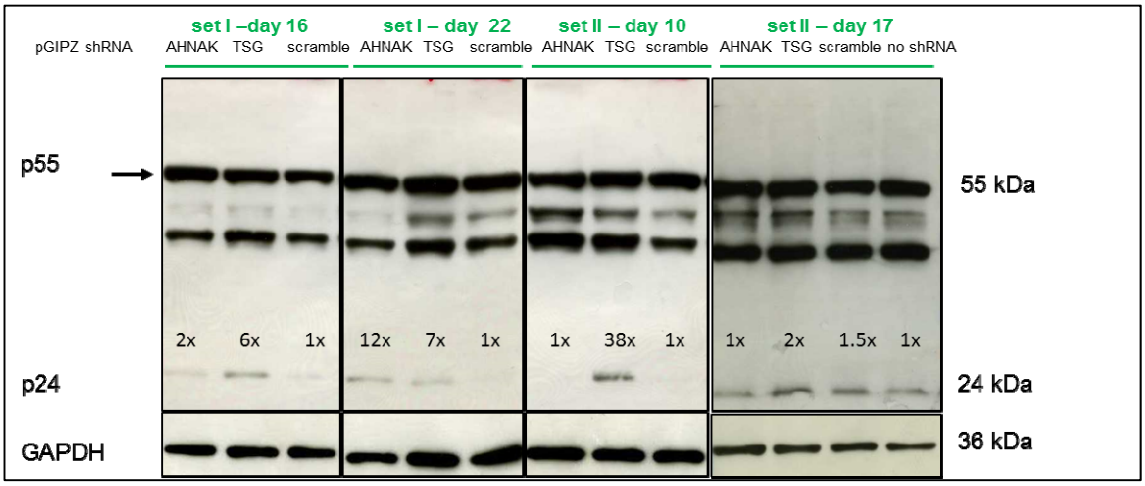


Figure 27: No considerable effect of shRNA mediated knock-down of AHNAK and TSG101 on Gag precursor and p24 levels in STAR-GIPZ cells; western blot of STAR-GIPZ-cell lysates in four sets of samples showed comparable levels of p24 in three out of four set; equivalent loading of total protein was verified by western blotting of GAPDH, 10 µg of total protein of cell lysates per lane.

5.3.2.2. 293T-GIPZ LV producer cell analysis

Following LV-GIPZ transduction of 293T cells and selection in puromycin they were transiently transfected to produce LVs followed by analysis of producer cells and vectors on (day 13 and 20 post-GIPZ transduction; set I). The GIPZ transduction and transient transfection procedures were repeated in 293T cells followed by analysis of cells and vectors 17 days after GIPZ transduction, representing set II. At each time point western blotting of AHNAK, ALIX, TSG101 and GAPDH was carried out to confirm that target host protein expression levels stayed reduced and were comparable before each vector production shows 293T-GIPZ-LV producer cell lysates of all three sets Figure 26.

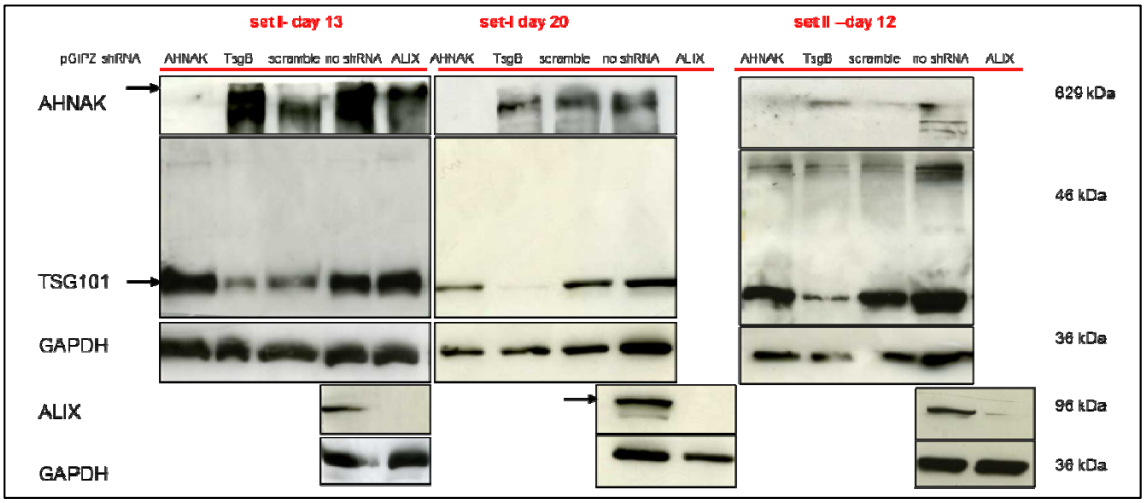


Figure 28: shRNA mediated knock-down of AHNAK, ALIX and TSG101 protein expression in 293T cells is stable over time. Assessment of knock-down efficiency of AHNAK, ALIX and TSG101 in 293T-GIPZ cell lysates in set I and II. 293T cells were transduced with GIPZ-LVs and selected in puromycin for at least 10 days and cell lysates analysed by western blotting of host cell proteins AHNAK, ALIX and TSG101; equivalent loading of total protein was verified by western blotting of GAPDH, 10 µg of total protein of cell lysates per lane loaded. Negative control were 293T cells without transduced GIPZ-LV ('no shRNA'). Protein molecular sizes are indicated.

To assess ALIX and AHNAK protein expression levels with or without transduction of STAR and 293T cells with GIPZ-LV, they were directly compared in western blots of cell lysates (Figure 27). This allowed comparison of the baseline protein expression before shRNA treatment as well as the comparison of shRNA knock-down efficiency in the two cell lines. Comparison of protein levels in STAR-GIPZ-ALIX and STAR-GIPZ-scramble shows that STAR and 293T cells express similar levels of AHNAK protein as well as similar levels of ALIX protein compared to protein levels in STAR-GIPZ-AHNAK and STAR-GIPZ-scramble, indicating that integration of vector components in STAR cells did not alter the expression of ALIX or AHNAK. A high level of knock-down efficiency seven days post-GIPZ transduction was achieved in both, STAR and 293T cells, in comparison to GIPZ-LV producers that expressed a shRNA targeting another host protein.

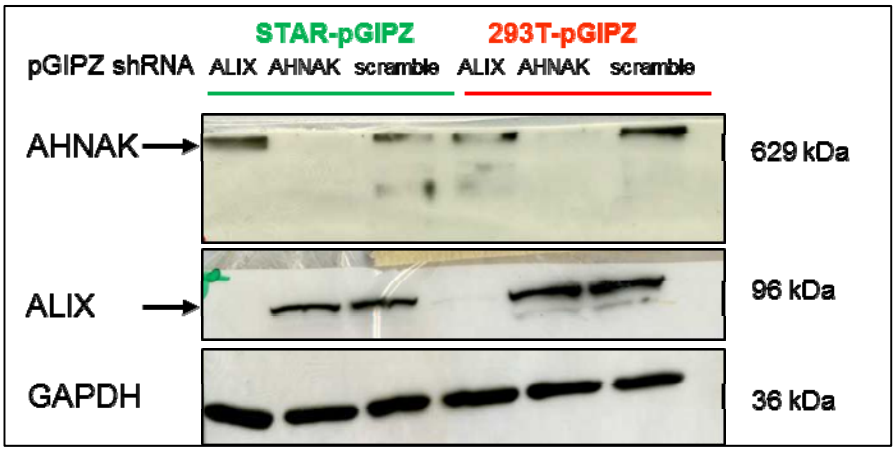


Figure 29: Knock-down efficiency of ALIX and AHNAK by shRNA is similar in STAR and 293T cells and baseline expression levels in both cell lines are comparable; western blotting allows direct comparison of protein knock-down efficiency in STAR-GIPZ and 293T-GIPZ 7 days post-GIPZ transduction. Equivalent loading of total protein was verified by western blotting of GAPDH (10 µg of total protein of cell lysates per lane). Negative control was 293T cells without transduced GIPZ-LV ('no shRNA'). Protein molecular seizes are indicated.

For transient LV production GIPZ-293Ts were co-transfected with p8.91 and pMDG as well as the vector genome plasmid pHV using Fugene transfection reagent. The same master mix was used for each GIPZ-293T sample (AHNAK, ALIX, TSG101, scramble and 'no shRNA' control) to avoid variation in transfection efficiencies. Transfection efficiency could not be determined by

counting the percentage of GFP positive cell using FACS analysis as 293T cells already express GFP from pGIPZ. 293T cells that were only transiently transfected to produce LVs, referred to as 'no shRNA control sample' were analysed by FACS for percentage of GFP positive cells to assess transfection efficiency. Transient transfection of 293T cells was carried out in three independent experiments and resulted on average in a relatively low transfection efficiency of 49.3%±6% (n=3) of GFP positive cells (Figure 28) in the 'no shRNA' sample. This means that only a fraction of 293T cells produced vector particles. An efficient transfection should result in >90% of transfected, GFP positive cells. All GIPZ-293T cells of one set were transfected with the same transfection mix resulting in comparable transfection efficiency for each sample and the percentage of transfected cells in the 'no shRNA' was used to estimate the percentage of transfected cells in GIPZ-293T cells. If the transfection efficiency in all samples was comparable and knock-down of the selected cellular host protein affected transfection efficiency, the proportion of transfected cells would be similar and any differences in production levels due to effects of protein knock-down.

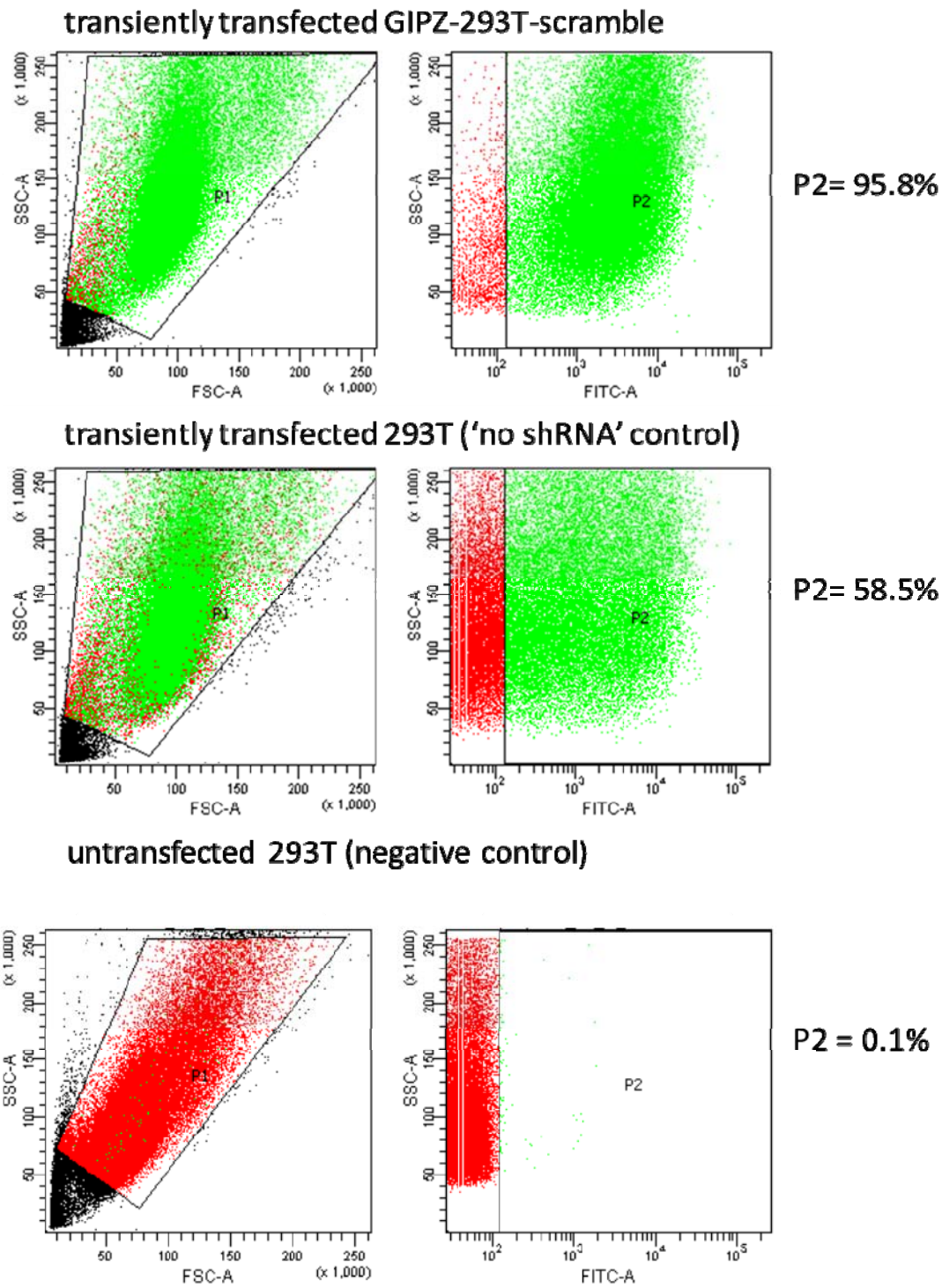


Figure 30: Assessment of transfection efficiency of 293T cells with GIPZ-LVs. FACS analysis of transiently transfected 293T cells expressing GIPZ-scramble (top) or no shRNA (bottom); representative example shows cells of one well of samples of set II day 12. top panel: after transduction with GIPZ-LV-scramble and transient transfection of plasmids for LV production 95.8 % cells express GFP, this is the sum of GFP expression from GIPZ-GFP expression cassette and transiently transfected vector genome, bottom panel: 293T cells that were only transiently transfected to produce LVs (without GIPZ-LV transduction) also referred to as 'no shRNA' control sample, contain 58.5% of transfected, GFP positive cells.

Host cell protein expression knock-down effects on vector production were also analysed in 293T-GIPZ cell lysates. Results for western blotting of viral protein p24 in producer cell lysate of each set of samples are shown in Figure 28. P24 levels are below the detection limit for all samples even when 30 µg total proteins were loaded (three times more compared to STAR-GIPZ cell lysates, Figure 25). Gag precursor levels of set I day 13 and day 20 are too weak to be detected. Low levels of vector proteins in producer cells are possibly due to low transfection efficiency of 293T-GIPZ (as described in section 5.3.2).

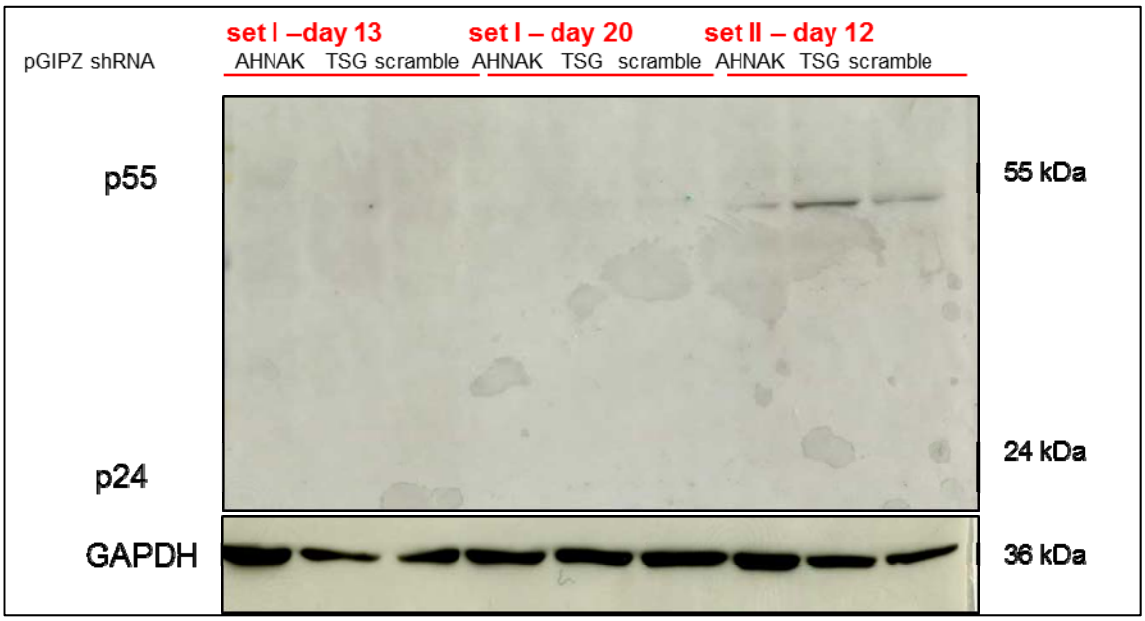
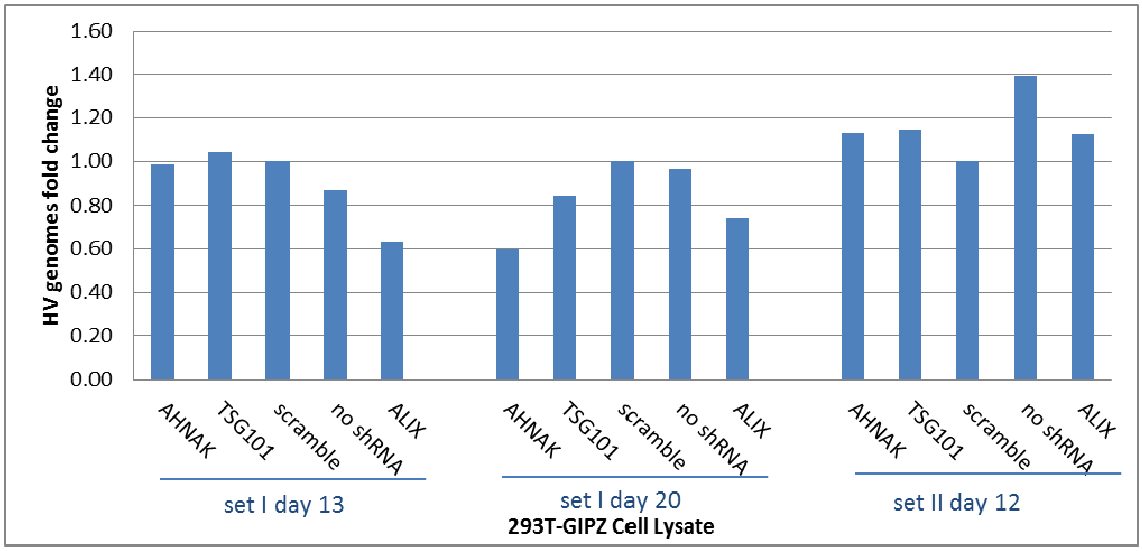


Figure 31: p24 levels in 293T-GIPZ cell lysates are below detection limit; gag precursor was detected only in samples of set II day 12 suggesting they contain highest levels of vector proteins and produced highest levels of LVs; equivalent loading of total protein was verified by western blotting of GAPDH (30 µg of total protein of cell lysates were loaded)

As in the experiments with STAR-GIPZ cells the fold change of vector genome RNA copies was quantified in 293T-GIPZ-AHNAK, 293T-GIPZ-TSG101 as well as 293T-GIPZ-ALIX producer cell lysates compared to 293T-GIPZ-scramble using Q-RT-PCR, normalised to human beta-actin copies per cell. Vector genome HV RNA copies in 293T-GIPZ-AHNAK and 293T-GIPZ-TSG101 producer cells analysed in three repeat-experiments were 0.63 to 1.39 fold of the vector

genome RNA copies in 293T-GIPZ scramble. 293T-GIPZ-ALIX producer cells of the first set contained 0.63 fold vector RNA copies compared to the scramble control. 293T-GIPZ-AHNAK producer cells of set I day 20 contained 0.6 fold vector RNA copies compared to the scramble control but these results did not repeat in subsequent sets (Figure 30). Overall there is no consistent difference to vector RNA copies in the scramble control.



					SFFV fold change in host cell KD samples compared to scramble
set	day	GIPZ	Ct beta actin	Ct SFFV	2^-deltadeltaCt
I	13	AHNAK	29.5	29.3	0.99
I	13	TSG101	30.5	30.3	1.04
I	13	scramble	29.4	29.2	1.00
I	13	no shRNA	31.6	31.6	0.87
I	13	ALIX	35.1	35.8	0.63
I	20	AHNAK	17.1	23.6	0.60
I	20	TSG101	17.3	23.4	0.84
I	20	scramble	16.9	22.7	1.00
I	20	no shRNA	16.7	22.5	0.97
I	20	ALIX	17.7	24.0	0.74
II	12	AHNAK	27.2	27.2	1.13
II	12	TSG101	27.8	27.8	1.15
II	12	scramble	26.9	27.0	1.00
II	12	no shRNA	26.1	25.7	1.39
II	12	ALIX	27.7	27.6	1.12

Figure 32: No consistent effect of shRNA mediated knock-down of AHNAK, ALIX and TSG101 on the number of vector genome RNA copies (HV) in 293T-GIPZ cells compared to 293T-GIPZ-scramble. Comparative quantification of HV RNA fold difference in cell lysates using Q-RT-PCR is shown. Results of vector genome RNA copy levels in three sets of samples; normalised to ß-actin copies per cell; negative control 293T-GIPZ-scramble and 'no shRNA' are comparable.

5.3.3. Analysis of Vectors Produced by GIPZ-Producer Cells

5.3.3.1. STAR-GIPZ Lentiviral Vector Analysis

Vector particle production after shRNA knock-down was analysed using several assays. First, the number of infectious vector particles was measured. 293T cells were transduced with vector from triplicate wells of samples of each set followed by FACS analysis of GFP expression of infected cells (Figure 30). Comparison of infectious titers of LVs produced by STAR-GIPZ-TSG101 showed that these are significantly lower than those of STAR-GIPZ-AHNAK produced vectors (p-value 0.0314, Figure 31). Statistical analysis was carried out using the two way analysis of variants, ANOVA. The two variants are 1) set of samples (set I and set II) and 2) days of analysis post-puromycin selection start. The average titer of vectors produced by STAR-GIPZ-TSG101 was $7.35x10^5$ TU/ml ($\pm 1.91x10^5$) and $1.14x10^6$ ($\pm 1.33x10^5$) produced by STAR-GIPZ-AHNAK. Overall the average titer difference of LVs produced by STAR-GIPZ-TSG101 and -AHNAK in four sets of samples is only 0.19 fold (Figure 31). Titers of vectors produced by STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 were between 0.7-2.1 fold and 0.6-1.0 fold the titer of the scramble control, respectively. STAR-GIPZ-ALIX produced vectors had an infectious titer of 1.1-2.4 fold compared to the scramble control (Figure 32). Overall no significant difference in infectious vector production compared to the scramble control was observed. Statistical analysis also showed that variation of titer over time is random and does not follow a specific trend (no significant interaction between set and time). No further analysis of vector production in STAR-GIPZ-ALIX cells was undertaken.

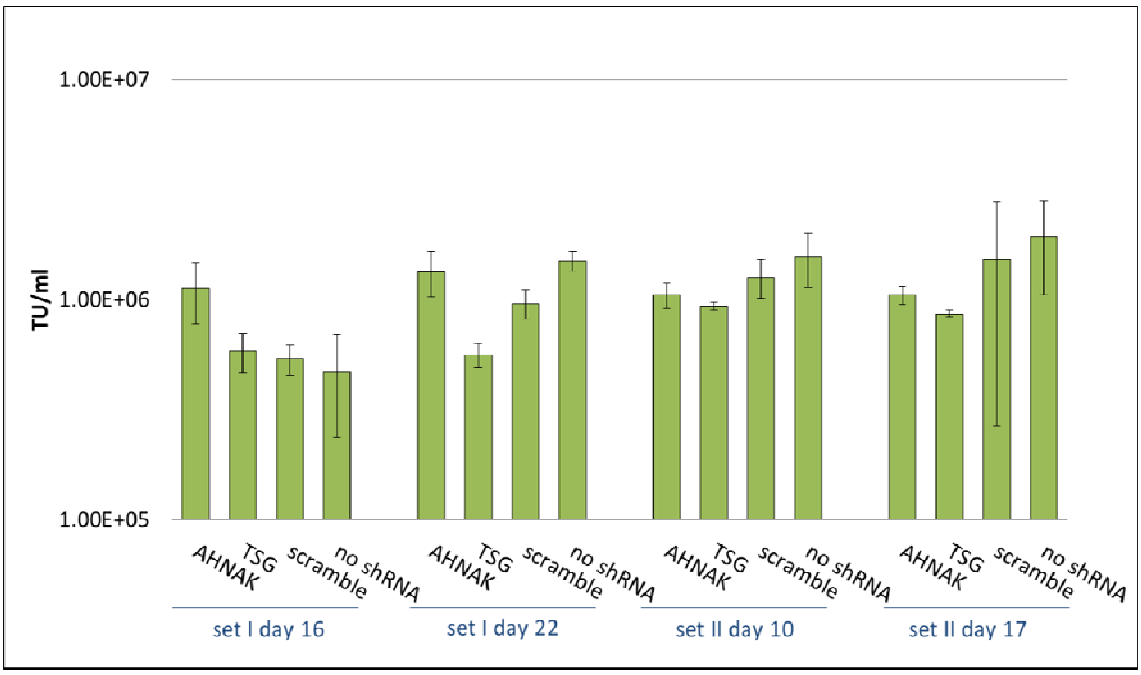


Figure 33: No significant effect of shRNA mediated knock-down of AHNAK and TSG101 on infectious particle production from STAR-GIPZ cells was seen. Infectious particle production from STAR-GIPZ cells from four sets of samples (transducing units/ml vector harvest) are shown and are the average of 3 wells/ sample per set. Bars on columns indicate SD.

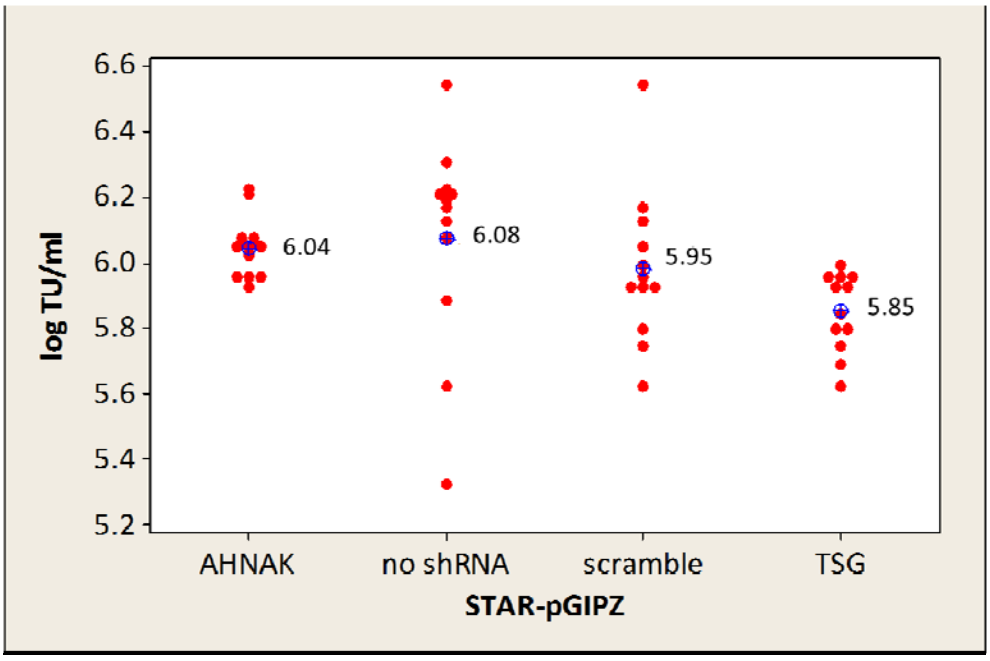


Figure 34: No significant effect of shRNA mediated knock-down of AHNAK and TSG101 on infectious particle production in STAR-GIPZ cells. Individual value plot of STAR-GIPZ produced vectors (log_{10} TU/ml). Mean values are displayed. Overall no significant difference in infectious titers compared to the scramble control (p-values >0.05). Variation of titer over time is random

and does not follow a specific trend (no significant interaction between set and time, p-value >0.05). Infectious titers of STAR-GIPZ AHNAK produced vectors were significantly higher compared to STAR-GIPZ-TSG101 produced vectors (p-value 0.0314) however are on average only 0.19 fold higher. Log TU/ml = log10 TU/ml.

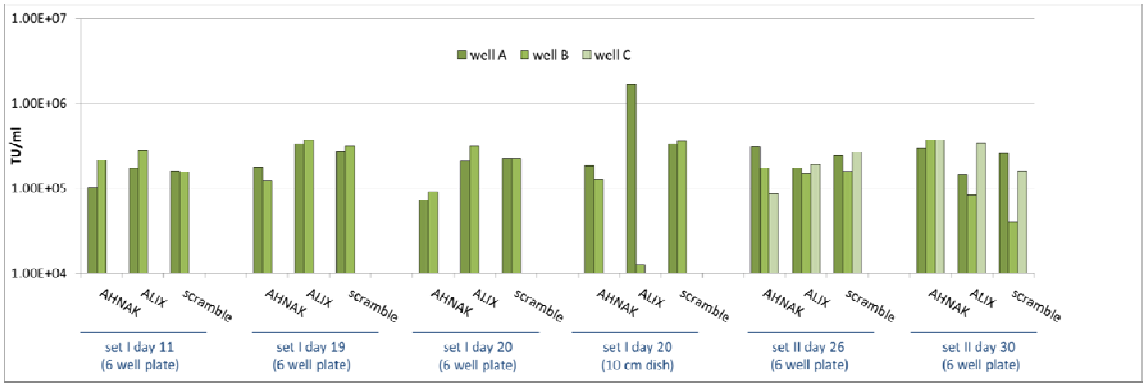


Figure 35: No significant effect of shRNA mediated knock-down of ALIX (and AHNAK) on infectious particle production from STAR-GIPZ cells was seen. Infectious particle number in STAR-GIPZ vectors (transducing units/ml vector harvest), duplicate or triplicate wells measured (A, B and C). Vector titers in ALIX knock-down produced samples on day 20 (10 cm) dishes are considered results of variable.

Further analysis of vector produced by STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 cells was carried out to assess if other aspects of vector production are affected. Q-RT-PCR showed that reduced levels of protein expression after knock-down, did not affect levels of viral genome RNA in producer cells (Figure 24). In order to investigate if knock-down of protein expression in producer cells influences the incorporation of viral genome RNA into vector particles, vectors were produced by the same number of STAR-GIPZ-AHNAK, STAR-GIPZ-TSG101 and STAR-GIPZ-scramble cells followed by measuring viral reverse transcribed RNA copy numbers in the vectors by Q-RT-PCR. Vectors of duplicate wells were analysed (duplicate wells: well A and B). The number of reverse transcribed vector RNA copies was calculated relative to a pHV plasmid standard at a known concentration. STAR-GIPZ-AHANK produced vectors were 0.6-1.8 fold of control RNA level (scramble) with one result of variable, 5.6 fold RNA copies compared to scramble control, in set I on day 16. STAR-GIPZ-TSG101 cells produced titers of 0.3-2.0 fold

compared to the control (Figure 34). Overall no significant changes in viral RNA levels were observed compared to scramble control or between different GIPZ producer vectors confirming results of quantification of RNA copies in producer cells where no significant difference between samples was detected either (Figure 24). These results suggest packaging of vector genome RNA is not affected by AHNAK or TSG101 knock-down in producer cells.

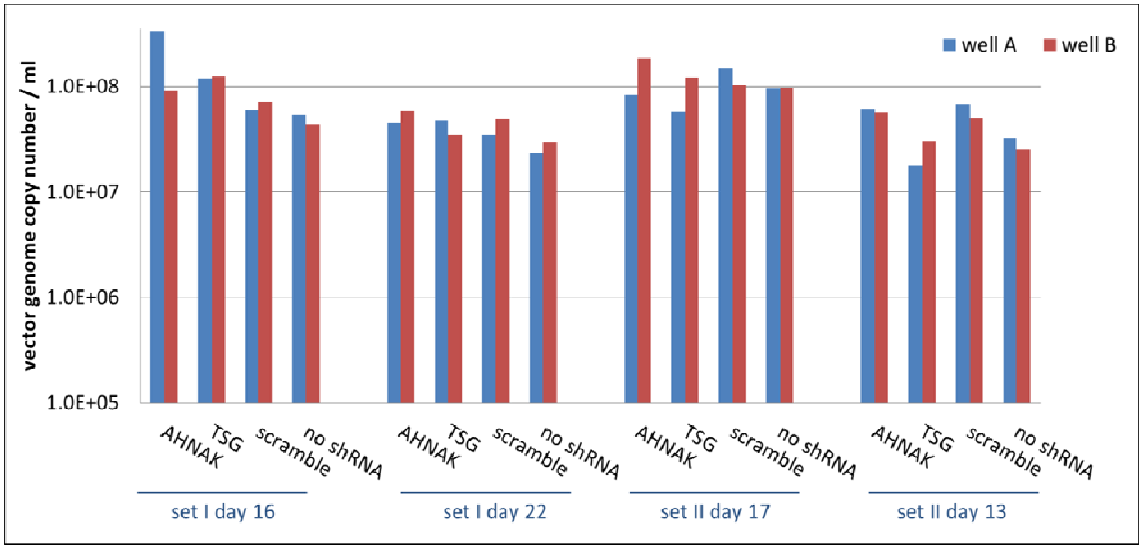


Figure 36: No effect on vector genome copy numbers in STAR-GIPZ vectors (reverse transcribed pHV RNA copies/ml vector harvest); four sets of samples were analysed. Quantification of reverse transcribed HV RNA copies in vector harvests and HV DNA plasmid standard using Q-RT-PCR is shown. Duplicate wells A and B for each sample; vector genome copies in vectors produced by STAR-GIPZ-AHNAK in well A of set I day 16 are 0.5 log higher compared to other samples, however this result did not repeat and was considered to be a result of variable.

P24 levels of vectors were measured by enzyme-linked immunosorbant assay (ELISA) in order to assess any differences in physical titre of vectors after knock-down of AHNAK or TSG101 expression. Quantities of p24 protein can be converted into physical vector particle numbers based on the assumption that one lentiviral vector particle contains 2000 molecules of p24 protein. The calculation is explained in Figure 34. The Particle number calculated based on p24 ELISA is later compared to particle numbers calculated based on assays measuring infectious titers and vector genome RNA copies in vectors.

2000 p24 molecules per lentiviral particle (LP)

1 p24 protein = 24 000 Dalton (Da)

1 Da = 1.66×10^{-24} grams (g)

➔ $2000 \times 24\,000 / (1.66 \times 10^{-24}) = 2.89 \times 10^{-5}$ pg p24 / LP
or 1 ng p24 = 3.33×10^{7} LP

Figure 37: Calculation of lentiviral particle number based on p24 protein measured in p24 ELISA. Based on the assumption that 1 ng p24 = 3.33×10^{7} LP.

P24 levels in vectors of one well per sample per set were measured. Vector p24 levels of STAR-GIPZ-TSG101 producers were 0.7-1.1 fold of p24 levels of vectors produced by the scramble control representing no significant difference. However STAR-GIPZ-AHNAK produced vectors contained significantly higher p24 levels compared to the scramble control (1.1-2.5 fold of scramble control) and compared to STAR-GIPZ-TSG101 produced vector p24 levels (p-value 0.0151 and 0.0035, respectively). P24 levels in LVs from STAR-GIPZ producer cells from four sets are grouped to illustrate this more clearly (Figure 36). Increased p24 levels in vectors from STAR-GIPZ-AHNAK producers were seen in three out of four sets of experiments but levels are comparable in set II day 17, which becomes clear when fold changes relative to STAR-GIPZ-scramble control are visualised (Figure 37). P24 levels in vectors from 'no shRNA' controls ranged between 110 and 180 ng/ml and in vectors from STAR-GIPZ-AHNAK between 218 to 144 ng/ml. Compared to the 'no shRNA' control, p24 levels in LVs produced by STAR-GIPZ-AHNAK are not significantly higher (p-value >0.05). These data and previously shown results of p24 levels in vector STAR producer cells (Figure 25) suggests that knock-down of AHNAK expression may only have a small or no significant effect on p24 expression and packaging into produced vectors.

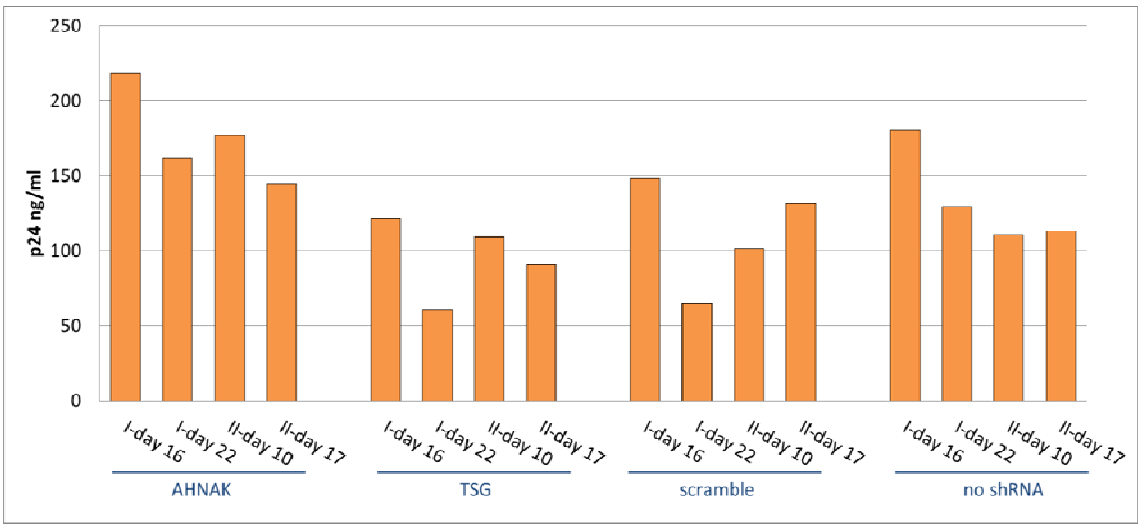


Figure 38: No consistent effect on viral protein p24 levels in STAR-GIPZ vectors; p24 in STAR-GIPZ vectors are shown grouped by STAR-GIPZ producer cells, p24 ng/ml of one well of vector harvests for each set was measured; in three out of four sets STAR-GIPZ-AHNAK vectors contained significantly higher numbers of vector particles compared to STAR-GIPZ-TSG101 and scramble produced vectors (p-value 0.0151 and 0.0035, respectively), this did not repeat in the fourth set of samples; no significant difference between STAR-GIPZ-AHNAK and 'no shRNA' control was measured.

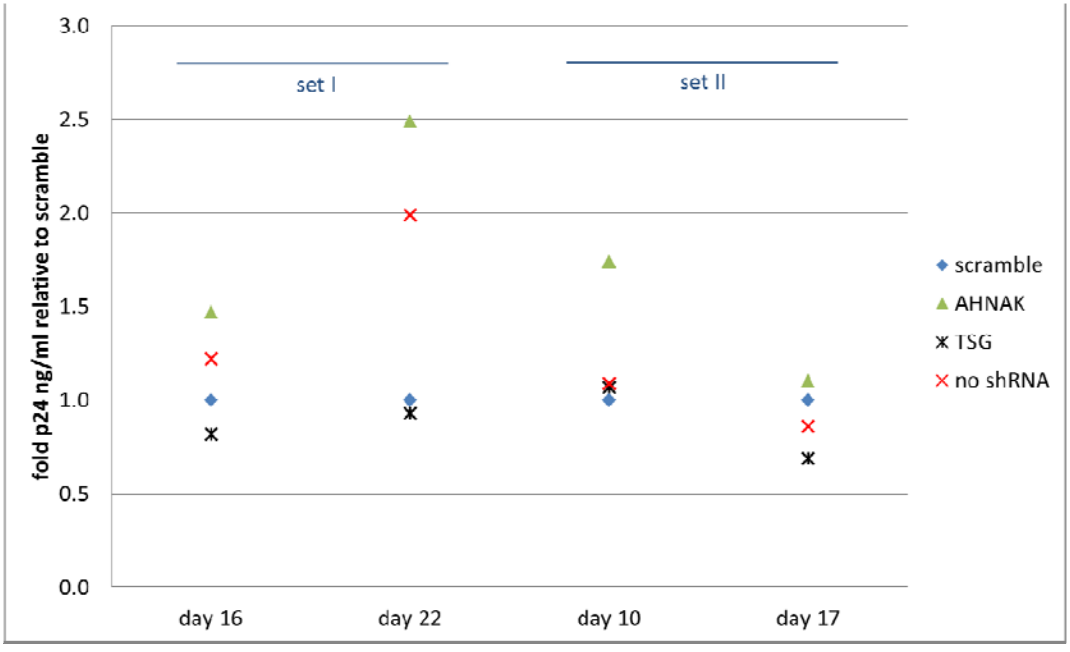


Figure 39: p24 levels in vectors of STAR-GIPZ-AHNAK, STAR-GIPZ-TSG101 and 'no shRNA' control relative to STAR-GIPZ-scramble (=1); fold p24 ng/ml of scramble p24 levels; in three out of four sets of samples p24 levels of STAR-GIPZ-AHNAK produced cells were at least 1.5 fold of p24 levels of scramble control; however p24 levels of all samples are comparable in the fourth set.

Protein p24 levels in vectors were further analysed by western blotting to assess the effect of AHNAK and TSG101 knock-down on vector production. Results for western blotting of vectors collected after selection in puromycin on day 16 and 22 (set I) as well as on day 10 and 17 (set II) are shown in Figure 38. No considerable difference between p24 values among LVs made by the different producer cells was detected. Based on densitometry readings of p24 signals in vectors STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 p24 levels varied between 0.8-1.9 fold and 0.6-1.3 fold of scramble p24 values, respectively (Figure 39). This method is not as sensitive as the p24 ELISA however can be used to confirm the relative p24 levels among different vector samples of the p24 ELISA. Western blotting showed that p24 levels of STAR-GIPZ-AHNAK vectors are comparable to the 'no shRNA' control (the 'no shRNA' control was only analysed for the fourth set of samples, set II day 17). Like in p24 ELISA, levels of p24 in STAR-GIPZ-AHNAK produced vectors are increased compared to STAR-GIPZ-scramble in three out of four sets of samples, however this difference is not significant. STAR-GIPZ-AHNAK and STAR-GIPZ-TSG101 produced p24 levels in vectors are comparable to each other, unlike results of p24 ELISA where p24 levels were higher in vectors produced from STAR-GIPZ-AHNAK than those from STAR-GIPZ-TSG101.

Interestingly the p55 Gag protein levels in vectors of STAR-GIPZ-TSG101 cells are 15 to 31 fold higher in three out of four experiments (Figure 40) compared to scramble control and STAR-GIPZ-AHNAK Gag precursor levels. This data suggests that Gag precursor processing to p24 is disturbed when TSG101 protein expression is reduced. This might only be a minor effect as elevated Gag precursor levels in vectors of STAR-GIPZ-TSG101 producer cells were not detected consistently in all sets of harvested vectors. Gag precursor levels in STAR-GIPZ-TSG101 are comparable to STAR-GIPZ-AHNAK and scramble in set II day 17 (Figure 38).

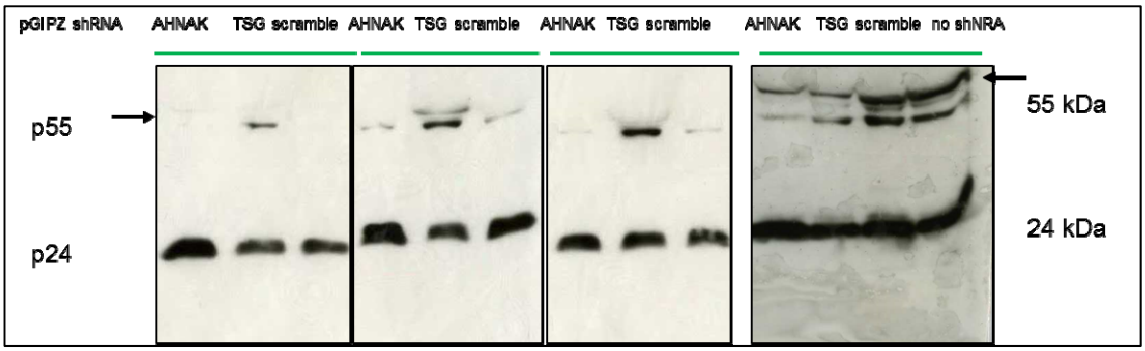


Figure 40: No effect of shRNA mediated knock-down of host protein AHNAK and TSG101 on p24 levels in STAR-GIPZ vectors; western blot using anti-p24 monoclonal antibody, proteins p24 and pre-cursor were detected; Gag precursor levels in STAR-GIPZ-TSG101 cells are significantly higher in three out of four sets of samples compared to STAR-GIPZ-AHNAK and scramble control (15 to 31 fold compared to scramble) however comparable in the fourth set; 45 µl of vector harvest per lane.

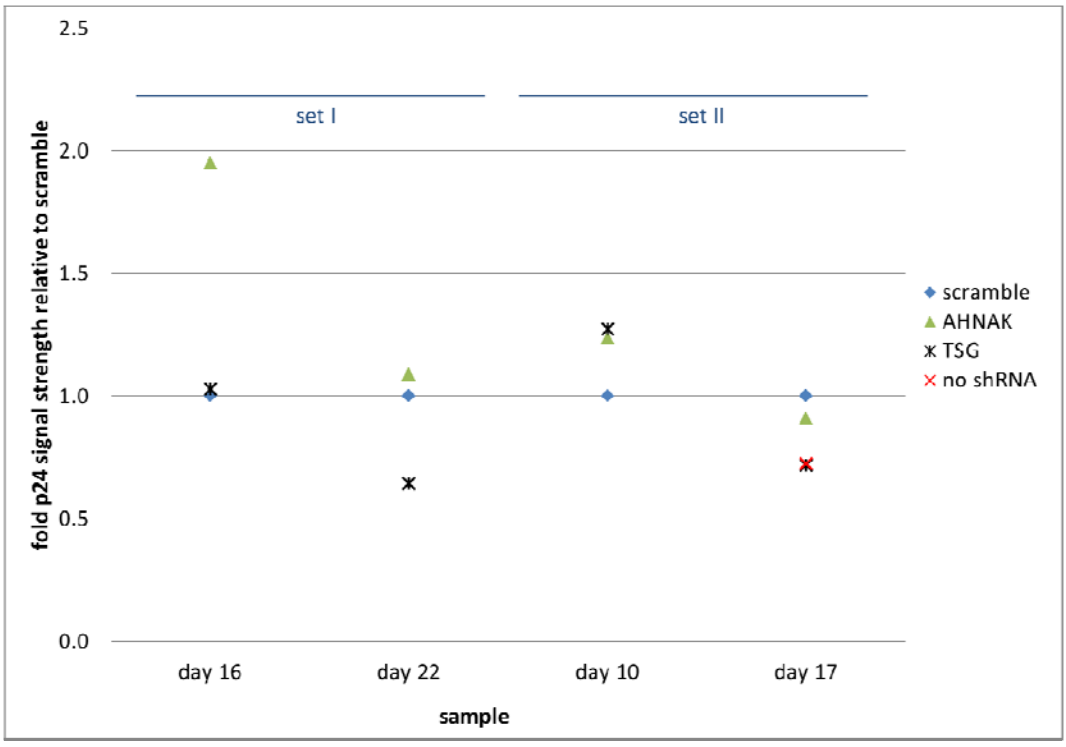


Figure 41: No effect of shRNA mediated knock-down of host protein AHNAK and TSG101 on p24 levels in STAR-GIPZ vectors; densitometry readings of p24 signal of western blots relative the scramble control (=1) is shown to visualise the minor differences between STAR-GIPZ samples and scramble control from Figure 38.

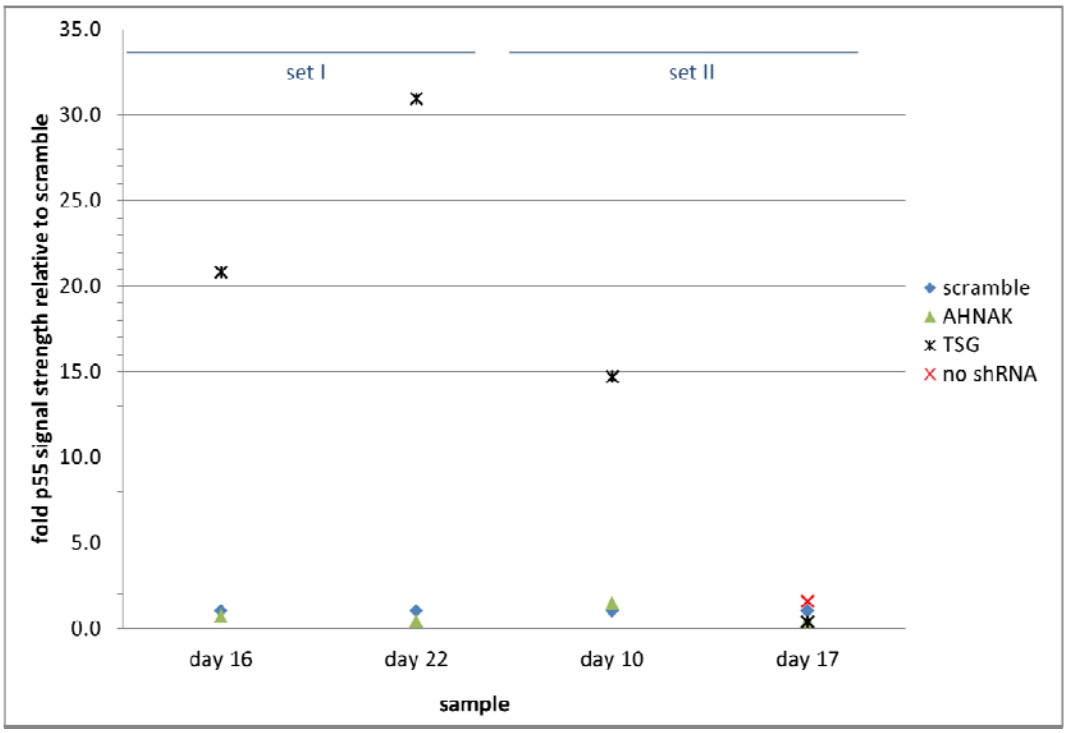


Figure 42: Effect of shRNA mediated knock-down of host proteins on Gag precursor levels in STAR-GIPZ vectors. Densitometry readings of Gag precursor signal of western blots relative to the scramble control (=1). Vectors from TSG101 knock-down STAR cells show a 15 to 31 fold increase of Gag precursor levels, however this effect was not observed consistently in all repeat experiments. STAR-GIPZ-AHNAK Gag precursor levels in vectors are comparable to those of controls.

5.3.3.2. 293T-GIPZ Lentiviral Vector Analysis

Vector particle production from 293T-GIPZ cells was also measured using several assays. After transduction of 293T cells with vectors produced by 293T-GIPZ cells, GFP positive cells were analysed for infectious particle production by FACS analysis. Supernatant of triplicate wells for samples of three repeat-experiments were measured. The number of infectious particles of samples from the first experiment showed about 10 fold lower transducing units per ml vector harvest compared to titers of the other two sets of experiments (Figure 41). This is most likely due to a less efficient transient transfection of cells in the first set of samples (set I day 13) with only 43% vector producing cells compared to about 55% in subsequent sets. Infectious

particles in vector harvests produced by 293T-GIPZ-TSG101 of set I are about five-fold higher compared to the control, but this was not seen in the following sets. In the other two sets of vector samples titers of 293T-GIPZ-AHNAK and 293T-GIPZ-TSG101 produced cells were 1.2 fold and between 0.9-2.1 fold of the titer of vectors produced by 293T-GIPZ-scramble, respectively. 293T-GIPZ-ALIX produced particles had an infectious titer of 0.6-0.8 fold compared to the control titer. Overall no significant difference in infectious titers compared to the scramble control was observed ($p>0.05$) (Figure 41).

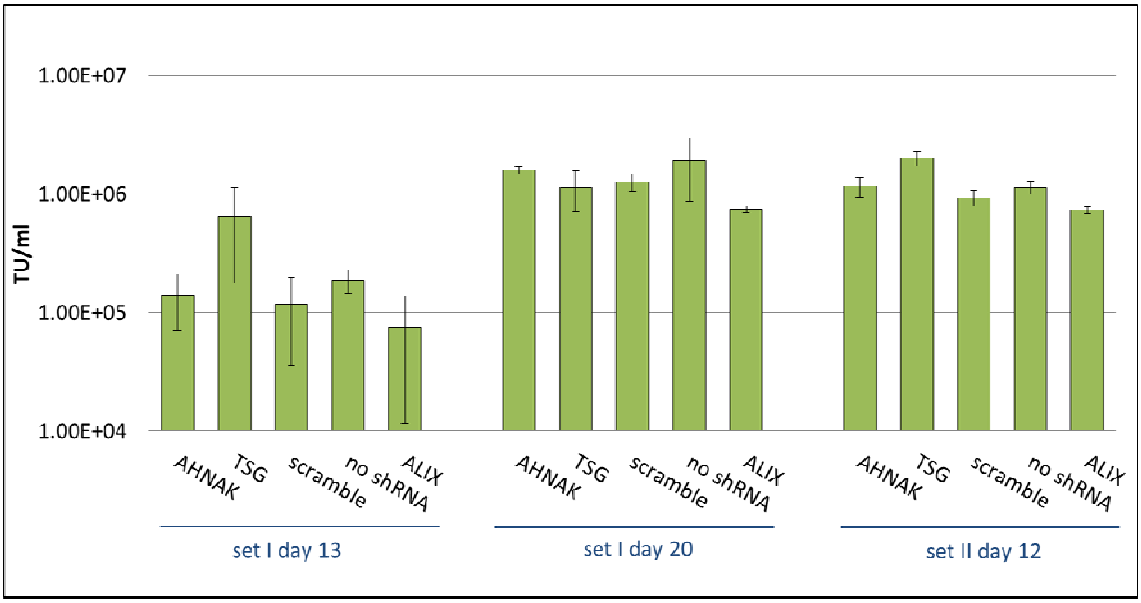


Figure 43: No significant effect of shRNA mediated knock-down of AHNAK, ALIX and TSG101 on infectious particle production in 293T-GIPZ cells ($p>0.05$). Infectious particle numbers in 293T-GIPZ vectors of three sets of samples (transducing units/ml vector harvest), shown is the average of 3 wells/sample per set.

Q-RT-PCR of vector genome RNA copies of 293T-GIPZ cells showed that production of vector genome RNA was unaffected by AHNAK, ALIX and TSG101 protein knock-down (Figure 30). The measurement of vector genome RNA copies in vectors from 293T-GIPZ producers was also carried out by Q-RT-PCR to investigate if packaging of viral RNA into particles was affected after gene knock-down. Vectors of duplicate wells were measured (well A and B). The number of reverse transcribed HV vector RNA copies was calculated using plasmid DNA pHV as the

standard. In some vector samples no vector RNA could be detected (for example well A, set I day 13 AHNAK, TSG101 and ALIX). This could be due to very low levels of vector genome RNA in these samples caused by low transfection efficiency (section 5.3.2). 293T-GIPZ-AHANK and 293T-GIPZ-TSG101 produced vectors were 0.4-1.9 and 0.9-1.6 fold of the scramble control titers. 293T-GIPZ-ALIX cells produced vectors with titers of 0.5-0.9 fold compared to the control. Overall no considerable titer changes were observed compared to scramble control or between different GIPZ producer vectors. In summary no effect on vector genome RNA packaging into vector particles could be detected.

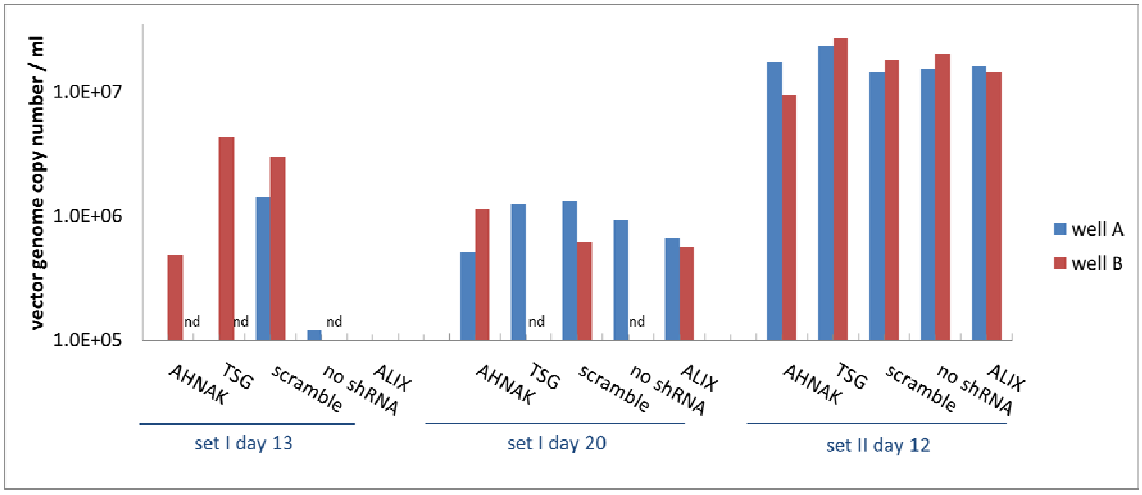


Figure 44: No considerable effect on vector genome copy numbers in 293T-GIPZ vectors (reverse transcribed pHV RNA copies/ml vector harvest); three sets of samples were analysed. Comparative quantification of reverse transcribed HV RNA copies in vector harvests and HV DNA plasmid standard using Q-RT-PCR is shown. Duplicate wells A and B for each sample; vector genome copy numbers of some of the vector samples of the first set of samples (set I day 13) were below the detection limit; however vector genome copy numbers in vectors of the other two sets were comparable to controls and showed no significant difference.

In order to assess any effect of host protein knock-down on vector capsid protein p24 its levels were measured in 293T-GIPZ vectors by enzyme-linked immunosorbant assay (ELISA). P24 levels can be converted into physical particle numbers by assuming 1 ng p24 equals $3.33x10^7$ LP physical particles. In a p24 ELISA vector harvests of one well per day were measured. P24

levels of vectors from 293T-GIPZ-AHNAK cells were 0.04 to 1.0 fold of the p24 levels in vectors produced by scramble. This included the result of the first set of samples when transfection efficiency was lowest resulting in about 10 fold lower levels of p24 compared to the other two sets confirming titration data of RNA copies and infectious particles of this set of samples compared to the other two sets. Differences in p24 levels of samples of set I day 20 and set II day 12 are not significant. 293T-GIPZ-TSG101 and 293T-GIPZ-ALIX produced vectors contained p24 levels 0.6 to 1.6 fold and 0.5 to 0.9 of p24 levels produced by the scramble control, respectively, representing no significant difference (Figure 43) ($p>0.05$).

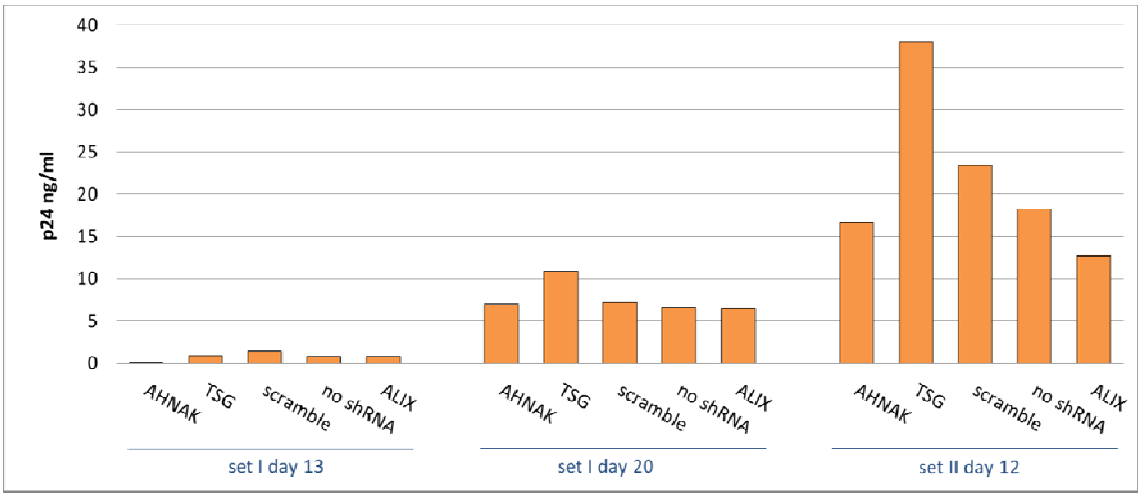


Figure 45: No significant effect on viral protein p24 levels in 293T-GIPZ vectors ($p>0.05$); p24 in 293T-GIPZ vectors are shown for three sets of samples, p24 ng/ml vector harvest of one well of vectors for each set was measured; the first set of samples contained about 10 fold lower levels of p24 compared to the other two sets, confirming results of previous assays presented.

Results for western blotting of p24 in vectors after AHNAK, ALIX or TSG101 knock-down in producer cells on different days are shown in Figure 44. P24 levels are below the detection limit for all samples. Hence based on this assay no conclusions about differences in p24 levels in vectors could be drawn. Gag levels of set I day 13 and day 20 are too weak to be quantified. Strongest Gag precursor signals could be recorded for samples of set II day 12, confirming

results of p24 ELISA and vector genome RNA copy number assay and titers of infectious particles, measuring highest vector particle levels in this set of samples.

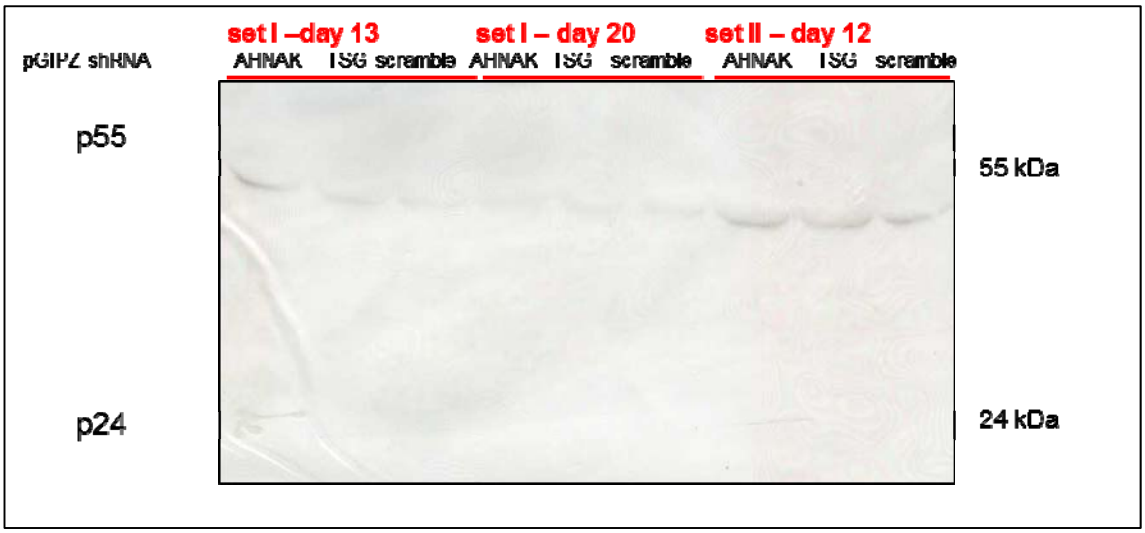
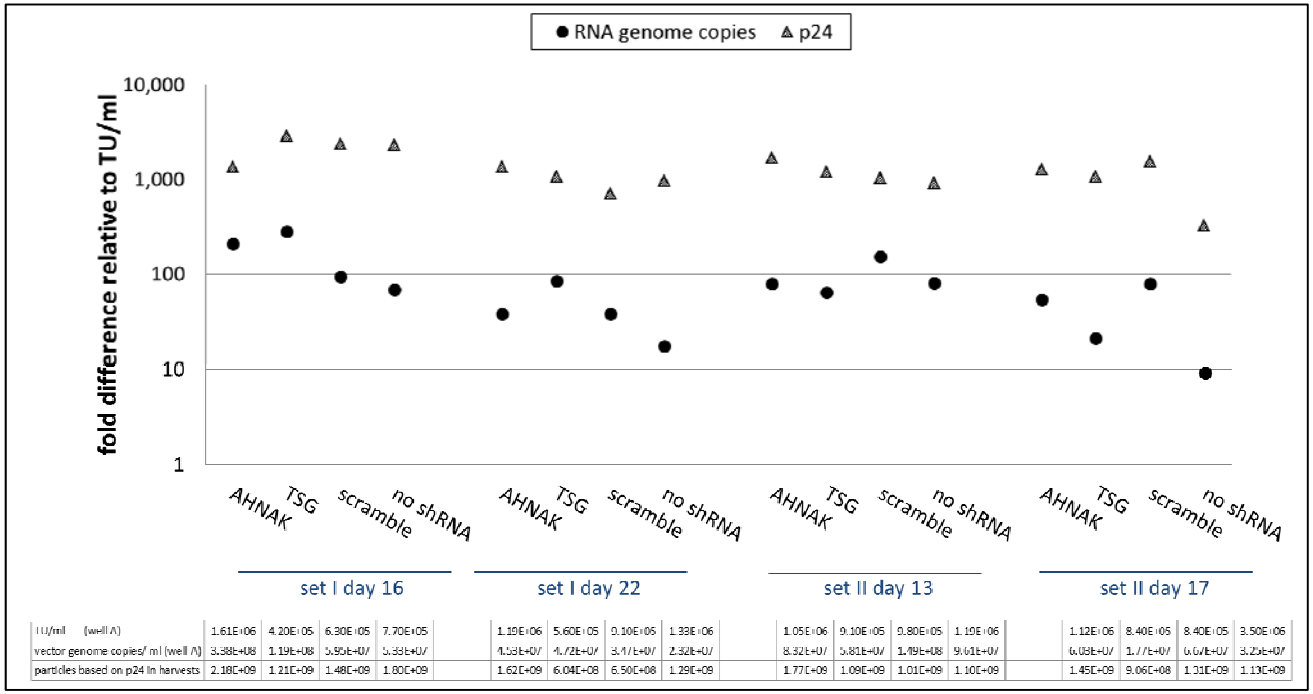


Figure 46: p24 levels in 293T-GIPZ vectors are below the detection limit, Gag precursor levels highest in the third set of samples confirming previous titration assays; western blot using anti-p24 monoclonal antibody is shown; 45 µl of vector harvest per lane

A general assessment of vector production levels in STAR-GIPZ and 293T-GIPZ based on various characteristics of harvested particles was undertaken in order to elucidate any potential differences between both producer cells. The ratio of infectious particles over physical particles based on vector genome RNA copy numbers and p24 levels (TU:RNA:p24) was calculated. Physical particle numbers are based on the assumption that one particle contains two copies of viral RNA genomes and 1 ng p24 equals 3.33×10^7 vector particles. The ratios of TU:RNA:p24 in STAR-GIPZ and 293T-GIPZ produced vectors show an interesting difference. RNA copies of STAR-GIPZ produced particles are on average 43 fold higher and corresponding physical particles based on p24 levels are on average 1000 fold higher compared to infectious particles. In contrast, vectors from 293T-GIPZ cells contain 3 fold higher physical particles based on RNA copies and 100 fold higher physical particles based on p24 levels compared to infectious particles meaning they are 10 times more infectious compared to STAR produced vectors (Figure 46 and 47). This has been shown in other experiments in this project as well as

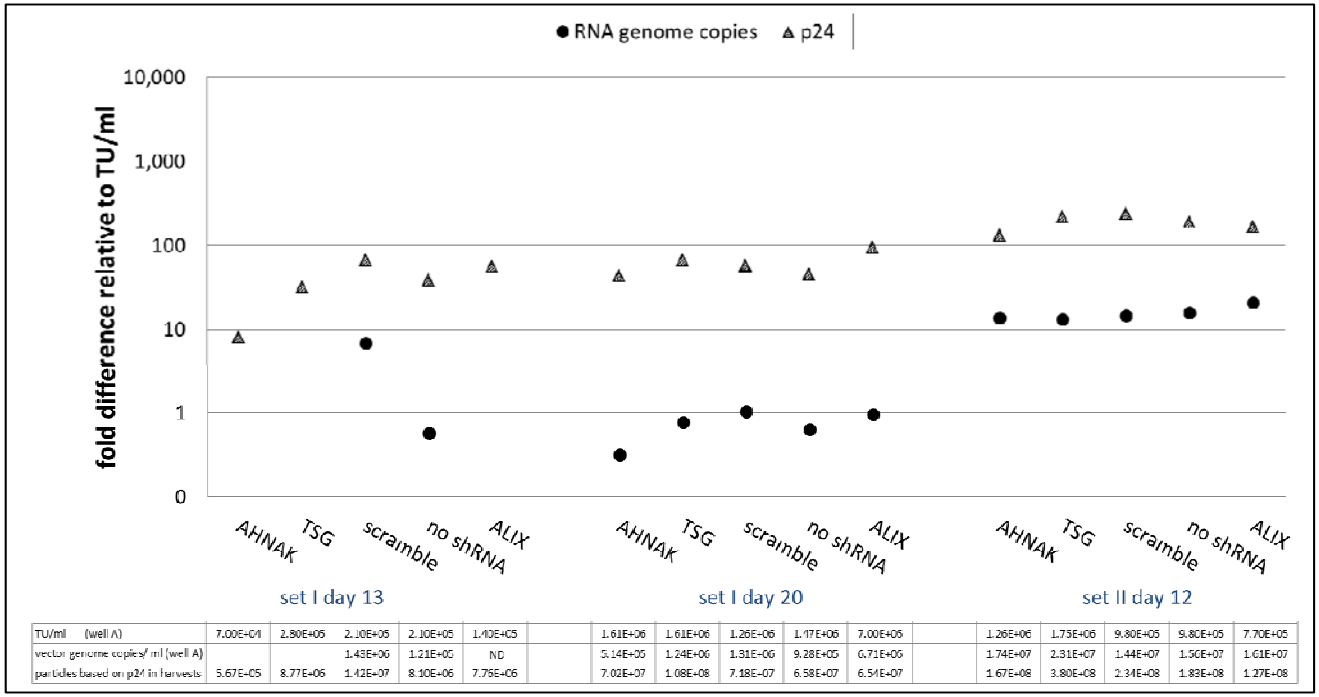
by others (Bell et al., 2010, Strang et al., 2004) and is most likely due to a lower infection efficiency of RDpro compared to VSV-G envelope.

Overall none of the titration assays showed that protein knock-down of AHNAK, TSG101 or ALIX had a significant, consistent impact in replicate sets of samples on vector particle production regarding infectious or physical particle numbers. In all assays the titer of the control 'no shRNA' varied between 0.5 to 2.0 fold of the titer of the scramble control with no significant difference in any of the experiments.



TU/ml (well A)	1.61E+06	4.20E+05	6.30E+05	7.70E+05		1.19E+06	5.60E+05	9.10E+05	1.33E+06		1.05E+06	9.10E+05	9.80E+05	1.10E+06		1.12E+06	8.40E+05	8.40E+05	3.50E+06
vector genome copies/ ml (well A)	3.38E+08	1.19E+08	3.95E+07	3.33E+07		4.53E+07	4.72E+07	3.47E+07	2.32E+07		8.32E+07	5.81E+07	1.49E+08	9.61E+07		6.03E+07	1.77E+07	6.67E+07	3.25E+07
particles based on p24 in harvests	2.18E+09	1.21E+09	1.48E+09	1.80E+09		1.62E+09	6.04E+08	6.50E+08	1.29E+09		1.77E+09	1.09E+09	1.01E+09	1.10E+09		1.45E+09	9.06E+08	1.31E+09	1.13E+09

Figure 47: STAR-GIPZ - ratio of RNA copies and p24 in vectors relative to infectious particles per ml of vector harvest.



TU/ml (well A)	7.00E+04	2.80E+05	2.10E+05	2.10E+05	1.40E+05		1.61E+06	1.61E+06	1.26E+06	1.47E+06	7.00E+05		1.26E+06	1.75E+06	9.80E+05	9.80E+05	7.70E+05			
vector genome copies/ ml (well A)			1.43E+06	1.21E+05	ND		5.14E+05	1.24E+06	1.31E+06	9.28E+05	6.71E+06		1.74E+07	2.31E+07	1.44E+07	1.50E+07	1.61E+07			
particles based on p24 in harvests	5.67E+05	8.77E+06	1.42E+07	8.10E+06	7.75E+06		7.02E+07	1.08E+08	7.18E+07	6.58E+07	6.54E+07		1.67E+08	3.80E+08	2.34E+08	1.83E+08	1.27E+08			

Figure 48: 293T-GIPZ - ratio of RNA copies and p24 in vectors relative to infectious particles per ml of vector harvests, samples, ND: not detected.

5.4. Discussion

In the experiments presented in this chapter host cell proteins that had been identified in size exclusion chromatography (SEC) purified lentiviral vectors by mass spectrometry (MS) are further analysed for their role in vector formation. Potentially MS detected host proteins might interact with viral proteins or RNA during vector assembly and budding from cells, which resulted in binding to or incorporation of the host proteins into secreted vector particles. The aim of the experiments presented in this chapter was to assess if knock-down of the selected MS-detected host proteins enhances or reduces LV production by quantifying vector particles.

Initially, several shRNA sequences targeting different parts of the mRNA of a selected protein were screened and the shRNA with the highest knock-down efficiency for each protein was chosen for further analysis. For the proteins AHNAK and ALIX a high level of knock-down of protein expression was achieved. The experiments performed showed no relation between the position of the shRNA target site within the mRNA sequence and the knock-down efficiency. The position of the binding site in the mRNA sequence of the shRNA sequence, that could efficiently knock-down protein expression, differed between the different proteins. Good knock-down efficiency was achieved by a shRNA targeting the 5'end of the mRNA of ALIX whereas high knock-down of AHNAK was achieved by a shRNA targeting the 3'end or middle part of the mRNA.

Selected proteins for further knock-down studies included AHNAK, ALIX and TSG101. AHNAK was identified by MS only in RD-pseudotyped LVs produced by STAR cells, hence studying the effect of AHNAK expression levels on vector production in the two different producer cell systems was of interest. ALIX was detected in all vector samples. ALIX is known to be an early-acting ESCRT factor (Katoh et al., 2003) and involved in HIV-1 budding (Votteler and Sundquist, 2013). TSG101 was not detected in SEC purified samples however is part of ESCRT-III host protein complex and recruits HIV-1 Gag via p6 to the budding site (Garrus et al., 2001,

VerPlank et al., 2001). Infectivity of HIV-1 secreted virions produced by TSG101 siRNA depleted 293T cells was reduced up to 10 fold (Martin-Serrano et al., 2003) thus it was thought to be suitable to be included in the study as a positive control. Significant knock-down of the candidate proteins EEF1A, ENO1 and MARCKSL1, identified by MS in all purified vector samples, was not successful and alternative shRNAs need to be screened. Hence these proteins were not further studied at this stage in the interest of time.

Assessment of vector infectivity is the most informative assay to characterise effects of protein-protein or protein-RNA interaction between host and vectors on vector formation. The production of infectious particles was not significantly affected after knocking down protein expression of AHNAK, ALIX or TSG101. Overall the infectivity of RDpro-pseudotyped vectors was 10 fold lower than that of VSV-G pseudotypes transiently produced (Figure 46 and 47) which has previously been shown to be caused by properties of the envelope protein used to pseudotype vectors. Ikeda et al. (2003) had reported a lower infectivity of stably produced gamma-retroviral pseudotyped vectors and observed that transiently produced vectors with gamma-retroviral envelopes were still equally less infectious compared to transiently produced VSV-G pseudotypes. Others produced LVs from STAR cells reported a 4-5 higher level of p24 ng/ml in vector harvests compared to VSV-G pseudotyped vectors when comparing LVs with similar transduction titer (Strang et al., 2004), confirming the results presented here. Other characteristics of vector particles were also analysed. No significant difference in RNA copy numbers was detected in vectors produced by knock-down cells. P24 levels of STAR-GIPZ produced vectors were unchanged compared to the negative control, whereas p24 levels in all 293T-GIPZ produced vector samples, including knock-down control scramble, were below the detection limit which may have been caused by low transfection efficiency of vector packaging plasmids.

The presence of an ALIX binding site in p6 of HIV-1 and its ability to recruit Gag to the vector assembly site (Strack et al., 2003) suggests that ALIX is relevant for HIV-1 budding from

infected cells. ALIX interacts with the C-terminal region the HIV-1 Gag p6 domain (Strack et al., 2003) that is required for HIV-1 particle release from the host cell membrane (Huang et al., 1995, Strack et al., 2003) and is classed as an early-acting ESCRT factor (Votteler and Sundquist, 2013). It was also found to be incorporated in HIV-1 virions produced by HeLa cells and HIV-1 virions with mutations in p6-Gag contained reduced levels of incorporated ALIX. Furthermore an interaction of ALIX and the p9 domain of Equine Immunodeficiency Virus (EIAV)-Gag was documented (Strack et al., 2003). Martin-Serrano et al. (2003) showed that ALIX interacts with proteins of the ESCRT-III complex hypothesising that ALIX recruits the ESCRT-III complex to the budding site. Based on these publications it was initially thought that knock-down of ALIX expression in producer cells would reduce LV production hence serve as a positive control when comparing to knock-down effects of AHNAK. However it was also shown that siRNA mediated depletion of ALIX in HIV-1 infected cells had only a minor effect on virus infectivity compared to relatively high titer reduction after knock-down of ALIX in EIAV infected cells. Consequently during EIAV budding ALIX is thought to be recruiting the ESCRT complex to the budding site (Martin-Serrano et al., 2003) which may explain the unchanged number of HIV-1 based infectious particles produced after ALIX knock-down in the experiments shown here and further analysis of knock-down effects of this protein was not undertaken.

TSG101 have been previously linked to retroviral budding from infected cells as a cellular partner for the p6 PTAP motif (Garrus et al., 2001, VerPlank et al., 2001). TSG101 has also been detected in MDM- and 293T-derived HIV-1 virions (Chertova et al., 2006, Saphire et al., 2006). Requirement of the interaction of TSG101 and p6 for HIV-1 release was demonstrated after mutation of the p6 PTAP motif lead to decreased TSG101 binding and infectivity of produced particles (Martin-Serrano et al., 2001) and supported by showing that infectivity after siRNA depletion of TSG101 in HIV-1 producing 293T cells is significantly reduced (Garrus et al., 2001, Martin-Serrano et al., 2003). Garrus et al. (2001) also showed that capsid (p24) and matrix (p17) levels are reduced in released virions after knock-down of TSG101.

Garrus et al. (2001) documented that TSG101 siRNA mediated depletion in 293T cells followed by transfection with proviral HIV-1, lead to reduced infectivity of produced virions of up to 14 fold, however in our experiments TSG101 knock-down in vector producing 293T cells or 293T derived STAR cells did not change the infectivity of the secreted vectors. In the results presented here p24 levels in 293T secreted vectors could not be detected even in control samples. In vectors produced by STAR-GIPZ-TSG101 no significant reduction could be seen in p24 levels compared to the scramble control. Martin-Serrano et al. (2003) did not specify the level of knock-down of TSG101 expression they achieved whereas Garrus et al. (2001) could only detect traces of TSG101 in cell lysates using an optimised knock-down system by transfecting siRNA twice at 24 hours intervals documented by western blotting after an additional 24 hours. Western blotting of TSG101 knock-down producer cells presented here showed that TSG101 expression was achieved efficiently, but does not result in complete protein depletion. Vector budding from producer cells might only require minimal levels of TSG101 and therefore need an even higher knock-down of TSG101. Differences in infectious or physical particle numbers could potentially be seen when TSG101 expression is even further reduced by using a higher dose of shRNA, by transducing producer cells with GIPZ-LVs at a higher MOI.

It is hypothesised that ALIX can bind p6-Gag and compensate for any functional loss regarding virus release when TSG101 expression levels are reduced (Fujii et al., 2007). Each protein binds a different HIV-1 p6 late domain (L-domain). TSG101, as part of ESCRT-I, is recruited to the budding site by binding the PTAP motif of HIV-1 p6 (Garrus et al., 2001) whereas ALIX binds the motif YPXL of p6 (Strack et al., 2003) as well as CHMP4 (Katoh et al., 2003) connecting ESCRT-III and HIV-1 Gag at the plasma membrane. Alternatively to the individual knock-down of either TSG101 or ALIX, TSG101 and ALIX could be knocked-down simultaneously to achieve reduced vector release from producer cells. Another approach would be to overexpress host cell proteins and study the effect on vector packaging. Usami et al. (2007) also showed that ALIX can bind via its C- terminal domain to TSG101. Overexpression of ALIX rescues HIV-1 release by

providing an alternative route for virus release after TSG101 interaction with PTAP is eliminated by PTAP domain mutation (Usami et al., 2007). Hence, knock-down of ALIX and TSG101 in the same producer cell might result in significant reduction of vector titer.

Whereas p24 levels among vectors were comparable in all STAR-GIPZ produced vectors, in consistence with the results of the infectivity and vector genome copy numbers, interestingly Gag-precursor levels in vectors, produced by STAR-GIPZ-TSG101 were 15 to 31 fold higher compared to any other samples. However, this effect might only be minor as it could not be seen in one out of four repeated experiments. As mentioned before, TSG101 binds Gag precursor, Pr55Gag, via the PTAP late domain of p6 of Gag (VerPlank et al., 2001). Furthermore, TSG101 depletion in 293T cells mediated by siRNA showed an increase of cytoplasmic levels of the Gag intermediate CA-p2 (CA-SP1) but not in released virions (Garrus et al., 2001). Nevertheless, raised levels of unprocessed Gag in vectors produced from STAR-TSG101 cells may be due to a vector maturation defect caused by TSG101 knock-down. A comparison to Gag precursor levels in 293T-GIPZ cells and vectors was not possible as protein levels were too low to be analysed by western blotting.

Knock-down of candidate proteins did not show effects on vector titers. In order to analyse viral protein expression from the producer cells after knock-down, viral protein levels within the producer cells were analysed. Producer cell analysis in all experiments showed that after knock-down the host cell proteins AHNAK, ALIX or TSG101 viral protein p24 levels or vector genome RNA copy numbers in producer cells were not considerably different compared to the scramble control. In transiently transfected 293T-GIPZ cells viral RNA and protein levels varied between different sets. In set I 5 to 10 fold lower levels of viral RNA copies were measured compared to producer cells of set II. Accordingly, vectors of this set also contained similarly reduced levels of infectious particles, RNA copies and p24 protein compared to the vectors of two subsequently produced set of samples and is potentially due to lower transfection efficiency when transiently transfected with vector plasmids (section 5.3.2).

The assay measuring transducing units of vector samples used in the study presented here quantifies vector particles that are secreted from the producer cells investigating only host-virus interactions in the late stages of the replication cycle, including transcription of viral genes to assembly and budding of LVs. The RNAi screen for host factors influencing HIV-1 infectivity presented by Zhou et al. (2008) investigated host proteins involved in early as well as late stages of HIV-1 production, the latter by investigating host factors that affected infectivity of secreted virus from knock-down producer cells. In the RNAi screen P4/R5 HeLa cells were used, which are modified HeLa cells that express CD4 and encode ß-galactosidase (Clavel and Charneau, 1994). The latter is expressed from *lacZ* under the control of a truncated HIV-1 LTR that upon infection of the cells with HXB2 virus is transactivated by Tat binding to the TAR sequence, therefore expression levels of ß-galactosidase are directly related to viral expression levels (Kimpton and Emerman, 1992). In a staged assay HeLa P4/R5 cells were transfected with siRNAs followed by HIV-1 infection (Zhou et al., 2008). Supernatant was harvested and added to HeLa P4/R5 cells that had not been treated with siRNA. They reported that virions produced by TSG101-depleted 293T cells were 20% less infectious compared to the negative control luciferase siRNA. The infectivity of 293T-GIPZ-TSG101 produced vectors in the experiments presented here was not reduced when compared to vectors produced by the scramble control. 293T-GIPZ-TSG101 produced vectors had a comparable infectivity to the scramble control, as seen when TU/ml levels were compared to p24 protein levels in 293T-GIPZ-TSG101 produced vectors and scramble control, which were comparable (Figure 46). Both vector populations have a similar amount of transducing particles per physical particle. This could be due to lower TSG101 knock-down levels in our experiments compared to the ones presented by Zhou et al. However, results of mRNA knock-down presented by Zhou et al. do not include mRNA knock-down levels of TSG101. Direct comparison of results is also difficult due to differences in several conditions in the experiment conducted by Zhou et al. and the conditions used in this study. Zhou et al. used HeLa cells (*versus* 293T or 293T based STAR cells in this study) and siRNAs were transiently transfected twice (*versus* transduction

with LVs delivering shRNA). Zhou et al. infected cells with HXB2 HIV-1 virus. An alignment of HXB2 *gag-pol* (GenBank K03455.1) and the sequence of p8.91 *gag-pol* in 293T transiently produced vectors showed they are 97% identical and STAR cells express codon-optimised *gag-pol* gene based on HXB2 *gag-pol.* However LVs produced here are pseudotyped with VSV-G or RDpro envelope and not the wild type glycoprotein envelope (gp120), potentially causing interaction with different host proteins during budding.

Apart from the study by Zhou et al. (2008), RNAi has also been used in two other genome wide screens to identify host cell factors in HIV-1 replication (Brass et al., 2008, König et al., 2008, Zhou et al., 2008). Brass et al. 2008 also investigated late stages of HIV-1 replication and infected P4/R5 indicator cells with virus produced by siRNA depleted cell, however in that and the other two screens neither AHNAK, ALIX, TSG101, MARCKSL1, EEF1A nor ENO1 were identified as a host factor in HIV-1 replication. Out of 842 genes identified in all three screens only 34 genes were identified as factors with potential roles in HIV replication in more than one screen (König et al., 2008, Brass et al., 2008, Zhou et al., 2008, Bushman et al., 2009). A total of 11 were previously reported in the NCBI HIV-1 interaction database and have been reported to play a role in the HIV-1 replication cycle. These include for instance T cell surface receptors cluster of differentiation 4 (CD4) and C-X-C chemokine receptor type 4 (CXCR4) which are HIV-1 target cell receptors providing host cell entry (Maddon et al., 1986, Feng et al., 1996) as well as the nuclear basket protein nucleoporin 153kDa (NUP153) involved in several steps of the HIV-1 life cycle for example the nuclear entry by binding HIV-1 integrase (Woodward et al., 2009). Only three genes were found to diminish HIV-1 replication in all three screens (mediator complex subunit 6 [MED6] and 7 [MED7] as well as RELA, also known as NfKB3 and only RELA has been linked to the NCBI HIV interaction database before. The little overlap in these siRNA screens was hypothesised to be due to variations in cell lines used, timing of sampling and filtration criteria (Bushman et al., 2009) and shows that the data presented in this project cannot be simply compared to these knock-down screens.

AHNAK could be a potential host factor during HIV-1 entry as its knock-down in $CD4^+$ T cells reduced HIV-1 ability to infect these cells (Wojcechowskyj et al., 2013), but apart from this study AHNAK has not been associated with retroviral replication or detected in 293T-produced LV particles. The more intriguing it was to investigate why of all analysed purified vector samples only STAR cell produced vector particles seem to be associated with this protein. If AHNAK is however expressed in higher amounts in STAR cells then this protein could potentially be secreted at higher levels into the cell culture medium compared to 293T cells and, if several proteins agglomerate could co-purify with vector particles instead of being directly associated to them. Semi-quantitative analysis of AHNAK protein however showed comparable levels in STAR and 293T cells. AHNAK has been localised to the plasma membrane in several cell types such as keratinocytes (Hashimoto et al., 1995) and in mouse embryonic fibroblasts (MEF) after binding to protein kinase C-α (PKC-α) (Lee et al., 2008). It is hypothesised that host cell proteins can be acquired by viruses through different mechanisms, such as positive selection caused by direct binding of the host protein to a viral protein in assembly and budding from the producer cell or through passive inclusion (for example tetraspannins) due to egress of the virus from sites expressing high levels of the protein (Ott, 2008). AHNAK could be packaged into vectors because of its localisation close to the vector assembly site at the plasma membrane.

Further analysis of host cell proteins needs to be undertaken to fully understand why they are identified in some but not all vector samples. Certain host cell proteins might be associated with vectors due to interaction with a viral protein or purely due to their abundant presence at the budding site and are not influencing production of LVs for that reason. It is unclear why AHNAK was only found in STAR derived RDpro pseudotyped vectors but not in VSV-G pseudotyped vector samples produced by transient production from 293T. However, absence of AHNAK in purified transiently produced vector samples needs to be confirmed by western blot analysis in order to confidently rule-out its association with VSV-G Env. AHNAK might not

have been detected in VSV-G pseudotyped transiently produced samples by MS due to high abundance of other proteins in the sample that may have masked detection of low levels of AHNAK.

Vector treatment with subtilisin has been used to eliminate host proteins that are not encapsulated within the particles (Segura et al., 2008) allowing the distinction between host proteins associated with the surface (or co-purified proteins) and vector incorporated proteins. Knock-down effects in other cell lines could also be investigated, for example T cell lines, as knock-down effects on vector production and infectivity could vary in different cell lines. Overexpression of host proteins in HIV-1 producer cells can lead to increased incorporation into virions due to its interaction with viral proteins or RNA. The latter has been shown for host cell protein Staufen. Higher incorporation levels of Staufen correlated with higher vector genome RNA levels in virions (Mouland et al., 2000).

The analysis of the selected vector associated host cell proteins AHNAK, ALIX and TSG101 did not identify host proteins whose expression levels are directly influencing vector production by improving or decreasing vector titers when reduced. Before further study of other host proteins such as EFF1A, ENO1 and MARCKSL1 is undertaken, knock-down efficiency needs to be improved. This could be either achieved by using a higher MOI during producer cell transduction with GIPZ-LVs. Alternative shRNA vectors could be used such as pSIREN, encoding a shRNA driven from the human U6 small nuclear promoter which is a RNA polymerase III promoter of one of the human RNA polymerases driving RNA transcription.

For complete elimination of protein expression knock-out methods can be used. During recent years naturally occurring DNA recognition codes have been engineered for example transcription activator-like effector nucleases (TALENs) (Boch et al., 2009) or clustered regularly interspaced short palindromic repeats (CRISRP) (Jinek et al., 2012). These systems can be used to induce double-strand breaks (DSB) in the DNA encoding the target protein. Through

cell induced repair of the DSB non-homologous end joining reconnects the DNA strands. This repair mechanism commonly induces errors leaving the expressed protein dysfunctional.

6. General Discussion

Lentiviral Vectors are being used for the treatment of a range of monogenic disorders, such as neurological diseases and haematopoietic disorders. Recently started clinical trials replace gamma-retroviral vectors with LVs due to their safer integration profile into the host cell genome, their ability to transduce non-proliferating cells as well as their larger transgene capacity. In addition, a recent publication of the outcome of a clinical trial shows that LVs are also suitable for cancer immunotherapy treating patients with acute lymphoblastic leukaemia (ALL). These developments show that the demand for LV supply is increasing. Large scale lentiviral vector production for example in the form of a stable producer cell line is required. Generally, vector production would benefit from a better understanding of vector particle formation within the producer cells, specifically interactions between host and viral proteins during particle assembly and budding. In this thesis cellular proteins associated with LVs produced by a transient and stable production system were identified by mass spectrometry (MS) and compared. Furthermore the effect of expression levels of selected proteins on vector titers was analysed.

Lentiviral vectors were prepared by the commonly used transient transfection method using VSV-G Env as well as by the packaging cell line STAR, continuously producing RDpro-pseudotyped vectors. It was noted that the transduction titer of crude samples was 10 fold higher in transiently compared to stably prepared samples, confirming lower titers of STAR-RDpro produced vectors compared to transiently produced VSV-G-vectors in previous publications (Ikeda et al., 2003, Strang et al., 2004, Bell et al., 2010). Increasing the concentration factor of RDpro-vectors six fold compared to VSV-G-vectors resulted in a comparable number of physical particles in purified samples produced by both methods for cellular protein analysis.

6.1. Discussion of Vector Purification

Size exclusion chromatography applying sephacryl 500-HR medium effectively purifies LV. It was shown that dominant proteins in purified samples are viral proteins and that viral vectors are separated from smaller, non-associated proteins. One limitation of this method was the relatively low recovery of infectious particles suggesting that TEN buffer may be affecting vector infectivity. An alternative elution buffer such as PBS or cell culture medium could increase recovery of infectious particles. Serum free Cell Gro® medium has been used as the buffer in size exclusion chromatography in vector preparation in a recently published clinical trial report (Aiuti et al., 2013). The use of TEN buffer is however suitable for the purpose of vector purification for a comparative MS analysis of vector particles.

The study presented here shows that SEC can be applied for small scale vector preparation in the context of *in vitro* research experiments and has been used in protocols in small scale clinical trials for vector purification (Levine et al., 2006). Anion-exchange chromatography (AEx), through resin or a membrane, has been used to purify HIV-1 derived (Schweizer and Merten, 2010, Merten et al., 2011, Aiuti et al., 2013, Biffi et al., 2013) and EIAV derived LVs (Stewart et al., 2010) in clinical trials and could be used to purify vector particles for a comparative analysis with SEC purified samples.

Purified VSV-G containing samples in this study could potentially contain tubulovesicular structures whose presence has been reported in sucrose-gradient concentrated LVs pseudotyped with VSV-G (Pichlmair et al., 2007). Proteins eluting in the void peak during SEC of VSV-G-only showed that this control sample, free of viral proteins, potentially contains protein complexes of similar size to vector particles, such as microvesicles. Their presence in purified VSV-G vector samples produced in this study would need to be confirmed for example by electron microscopy (Pichlmair et al., 2007). An additional control sample of VSV-G vectors treated with a protease, such as subtilisin, could be included in the mass spectrometry analysis to identify cellular proteins inside vector particles.

Lentiviral vectors purification from producer cell derived proteins in cell supernatants can reduce cytotoxicity and increases transduction efficiency. Current protocols used for vector production in clinical trials include steps that eliminate contaminants including DNA, derived from transfected DNA plasmids or lysed producer cells, using benzonase (Sastry et al., 2004) as well as purification steps such as anion-exchange chromatography for the removal of serum proteins derived from cell culture medium or proteins derived from producer cells (Yamada et al., 2003). However these methods are not able to remove potentially present impurities that have similar physical properties to those of vector particles. Anion-exchange chromatography is used to separate vector particles from impurities based on the negative charge of vectors binding to the positively charged chromatographic supports such as diethylaminoethanol (DEAE) anion-exchangers (Merten et al., 2011, Aiuti et al., 2013, Biffi et al., 2013), suggesting that protein complexes of similar charge could co-purify with vectors during AExc. Electron microscopy can be performed (Transfiguracion et al., 2003, Pichlmair et al., 2007) to visualise the presence of potential microvesicles in purified vectors. To analyse if specific proteins are located on the surface or inside purified vector particles, immunogold staining of vectors can be used (Segura et al., 2008).

6.2. Discussion of MS results and Identified Protein Functions

Vector samples of transiently produced VSV-G-vectors and stably-produced RD-vectors were prepared for mass spectrometry (MS) analysis. The protein composition of comparable numbers of viral particles based on p24 levels was analysed in several sets of each vector sample as well as controls as described in chapter 4.

The aim of this study was the identification of any true differences in vector associated host proteins among vector samples which can only be ensured if possible co-purification of comparable sized proteins or protein aggregates is excluded as they could be identified as false-positive vector associated proteins. The most striking difference was that STAR-RDpro-produced vectors contained substantially fewer protein species than transiently prepared VSV-G pseudotypes, hence in comparison stably produced RDpro pseudotyped vectors are of higher purity. Less different cellular proteins than in VSV-G pseudotyped vectors were also detected in envelope free Gag-Pol samples similar to RDpro-pseudotyped stably produced samples. In the samples analysed here detected cellular proteins that are exclusive to SEC-purified samples containing VSV-G protein could be derived from potential VSV-G vesicles of similar size to vectors that could have co-eluted in the SEC void peak fraction as they have been shown to be present in VSV-G pseudotyped lentiviral vectors (Pichlmair et al., 2007). The results presented here suggest that vesicles closely related in size to vectors are present at much lower levels or not at all in SEC-purified vectors pseudotyped with the gammaretroviral envelope RDpro. Proteins of VSV-G vesicles could be eliminated by protease digest for example using subtilisin, of vectors which has been used for protein digestion on the surface but not inside the HIV-1 virions (Ott et al., 1995) followed by sucrose-gradient centrifugation microvesicles in vector particle samples were reduced (Ott et al., 1996). Comparing the protein composition of transiently produced protease-treated VSV-G vector samples with untreated vectors could discriminate co-purifying, VSV-G vesicle derived, cellular proteins or proteins associated to the vector surface from vector incorporated proteins. In order to analyse if

differences in cellular protein composition are due to differences due to the packaging system used or due to the envelope protein of vectors MS analysis of SEC-purified vectors pseudotyped with other envelope proteins such as amphotropic murine leukaemia virus (MLV-A) Env, modified gibbon ape leukaemia virus Env (GALV+) or Rabies virus glycoprotein (RVG) would be interesting. Finally, a complete comparison would include the MS analysis of transiently produced RDpro-pseudotypes.

The predominant functions of MS-identified cellular proteins are cytoskeleton organisation, vesicular transport but also cell death. Some of the identified proteins are known to assist in trafficking of viral proteins to assembly site for example actin, tubulin as well as budding from cell for example ALIX, clathrin. This shows that other proteins we detected, such as MARCKSL1 and ENO1, that have not been shown to function in HIV-1 assembly, could be associated with purified vectors due to a role in vector assembly and budding. The presence of proteins functioning in necrosis and apoptosis could be due to induced stress of vector production in the cells. Many of the proteins, such as proteins of the cytoskeleton or cytoskeleton regulators as well as proteins associated with vesicular trafficking and transmembrane proteins, that were identified in our study have been found in mass spectrometry analysis of various sucrose gradient purified enveloped viruses in other publications (Chertova et al., 2006, Zhu et al., 2005, Varnum et al., 2004, Johannsen et al., 2004, Shaw et al., 2008, Moerdyk-Schauwecker et al., 2009) suggesting similarities in assembly and budding processes of different enveloped viruses resulting in the association of similar cellular proteins, for example secretory pathway trafficking of envelope proteins. Overlapping MS-results of our study with a study on HIV-1 virions produced in a different cell type, MDMs, compared to 293T cells used here, suggests that at least a subset of these host cell proteins is not randomly incorporated (Chertova et al., 2006). Analysis of vector titers after reducing gene expression levels of these proteins could elucidate which of them affect vector formation.

Generally, results of this study show that MS analysis is a suitable method for characterisation of viral vectors and could be applied for the comparison of other vector designs.

6.3. Discussion of Protein Knock-Down

Specific proteins that had been identified by mass spectrometry were selected for further analysis by knock down of gene expression in producer cells. Out of the selected proteins efficient knock-down was achieved for AHNAK, ALIX and TSG101 expression in STAR-RDpro as well as 293T cells. Analysis of vectors produced in cells after knock-down of expression of ALIX, AHNAK or TSG101 showed that infectious titers as well as physical titers of vectors were comparable to cells in which gene expression was not knocked-down; hence reduced protein levels did not have a measurable effect on LV production. One explanation could be that these cellular proteins do not influence LV assembly or budding in producer cells. Depletion of ALIX in HIV-1 infected cells by siRNA resulted only in a titer reduction of approximately 5 fold compared to about 15 fold reduction in titer after knock-down of ALIX in EIAV infected cells (Martin-Serrano et al., 2003). ALIX may only play a minor role in HIV-1 budding and may be more important in EIAV budding.

Infectivity of HIV-1 was significantly reduced after siRNA depletion of TSG101 in HIV-1 producing 293T cells (Garrus et al., 2001, Martin-Serrano et al., 2003). Martin-Serrano et al. (2003) did not specify the level of knock-down of TSG101 expression they achieved when using siRNA. Garrus et al. (2001), however, optimised their knock-down system by transfecting siRNA twice at 24 hours intervals to achieve almost complete depletion of protein expression as they could only detect traces of TSG101 proteins in cell lysates by Western blotting after an additional 24 hours. Increasing the efficiency of TSG101 knock-down in 293T or STAR-RDpro cells used in our study could potentially result in a better knock-down reducing residual TSG101 mRNA levels. For complete elimination of a specific protein, gene expression can be knocked out by using sequence specific nucleases such as transcription activator-like effector nucleases (TALENs) (Boch et al., 2009), clustered regularly interspaced short palindromic repeats (CRISRP) (Jinek et al., 2012) or zinc finger nucleases (ZFNs) which can be delivered by

viral vectors or directly by transfection (Gaj et al., 2012), disrupting gene expression or resulting in dysfunctional protein expression.

Both Garrus (et al. 2001) and Martin-Serrano (et al. 2003) were studying knock-down effects in wild type HIV-1 with the gp120 envelope protein. Knock-down of ALIX and TSG101 could potentially affect assembly and budding of VSV-G and RDpro-pseudotypes differently compared to wild type HIV-1 virions with the gp120 envelope protein. Both, ALIX and TSG101 bind HIV-1 Gag L-domains suggesting that they recruit the viral protein to the assembly side suggesting that binding of ALIX to p6-Gag compensates for any functional loss regarding virus release when TSG101 expression levels are reduced. This could also be the case vice versa. TSG101 could compensate for ALIX in ALIX knock-down cells vector particle budding.

AHNAK could be a potential host factor during HIV-1 entry as its knock-down in $CD4^+$ T cells reduced HIV-1 ability to infect these cells (Wojcechowskyj et al., 2013), but apart from this study AHNAK has not been associated with retroviral replication or detected in 293T-produced LV particles. AHNAK could associate with vector particles due to its location at the plasma membrane, hence close location to the assembly site. Why AHNAK was only detected in STAR cells is unknown, as cell lines, 293T as well as STAR cells express similar levels of AHNAK. It could specifically associate with RDpro Env. To investigate this possibility, an additional vector sample, transiently produced RDpro-pseudotyped vectors, would need to be included in this study.

RNAi screens in producer cells can be carried out for analysing the influence of cellular protein expression on vector production as presented by Zhou et al. (2008) for HIV-1 viruses. These RNAi screens could be applied for investigation of the involvement of cellular proteins in vector assembly and budding by investigating the transduction ability of vectors produced in knock-down producer cells.

6.4. Concluding Remarks

In conclusion, this project has characterised the cellular protein composition of SEC-purified LVs stably produced from the STAR cells pseudotyped with RDpro envelope and from transiently produced from 293T cells pseudotyped with VSV-G. Vector particles from both production systems were purified and directly compared regarding their cellular protein composition. Stably produced RDpro-pseudotyped vectors contained substantially less cellular proteins than transiently produced VSV-G pseudotyped vectors. Overall, mass spectrometry detected proteins include proteins, such as actin, tubulin, ALIX and clathrin that are known to function in HIV-1 virion assembly and budding. The characterisation of the identified cellular proteins, potentially associated with vectors through their interaction with viral proteins, provides a starting point for further studies on vector particle formation within the producer cells, specifically at the stages of vector assembly and budding.

7. Future Work

The following aspects can be included in future experiments:

Vector Samples. A transiently produced RDpro-pseudotyped vector sample should be included to distinguish proteins associated with the envelope protein from proteins that are associated with the viral core. This additional sample would also allow determining if, like in stably produced RDpro-pseudotypes, less host proteins are associated with transiently RDpro-pseudotypes compared to VSV-G pseudotyped vector samples. Further control samples, such as harvests from untransfected 293T cells or from cells transfected with DNA plasmids devoid of HIV-1 or VSV-G sequences, can be included. To prevent contamination with bovine proteins, vectors could be produced in serum free medium; replacing complete DMEM 24 hours post-transfection (see section 2.3.1).

Vector Purification. For the removal of non-viral membrane vesicles and other potential impurities closely related to the size and shape of functional LVs, rate zonal ultracentrifugation using iodixanol, a density gradient medium, could be employed (Segura et al., 2006a). An alternative or additional vector purification method can be tested, such as anion-exchange chromatography (AExc) which has been used to purify vector samples for their use in clinical trials. AExc could then be followed by SEC to remove smaller proteins that are of similar charge to vector particles. During SEC, an alternative buffer could be used such as PBS or serum free cell culture medium, similar to the gel filtration step in vector preparation for clinical trials (Aiuti et al., 2013).

Vector Sample Protein Analysis. For a complete analysis by MS, the same conditions of data acquisition should be used for all vector samples, such as the rejection method of transiently produced samples as well as setting the threshold to the same minimum of peptides that is required for the confident detection of a protein (see section 2.3.3.2.). This would potentially detect low abundance peptides in VSV-G pseudotypes that were not detected using the

normal acquisition method, allowing a more accurate estimation of the number of proteins detected in VSV-G-pseudotypes compared to RDpro-pseudotypes.

Purified vector particles could be visualised by electron microscopy of crude and purified vector particles in order to assess if any contaminating vesicles are present (Transfiguracion et al., 2003, Pichlmair et al., 2007). Subtilisin treatment of vector samples has been used to eliminate host proteins that are on the surface of vector particles (or co-purified along vector samples) (Segura et al., 2008) allowing the distinction between vector encapsulated host proteins. Alternatively the localisation of a host protein either on the surface or inside of purified vector particles can be investigated using immunogold staining of vectors (Segura et al., 2008). A co-immunoprecipitation assay of cell lysates might allow identifying interactions of host and viral proteins.

Analysis of Effect of Expression Levels of Specific Host Proteins on Vector Titers. TSG101 and ALIX might need to be knocked-down simultaneously to achieve reduced vector release from producer cells as they are binding different L-domains, PTAP (Garrus et al., 2001) and YPXL (Strack et al., 2003), respectively. An increase in knock-down efficiency might be necessary which could be achieved by using a higher dose of shRNA, by transducing producer cells with GIPZ-LVs at a higher MOI. In addition, for increased knock-down efficiency, an alternative shRNA vector could be tested such as pSIREN, encoding a shRNA driven from a promoter of a human RNA polymerase. Alternatively knock-out techniques could be used such as CRIPRS or TALENs (see section 5.4.).

The knock-down effect on vector production of other proteins that are common to all vector samples and are less well characterised in terms of viral vector assembly, should be investigated next, including EEF1A, MARCKSL1 and ENO1 (see Table 15). On the other hand overexpression of candidate host cell proteins might allow the identification of proteins that increase vector titers and hence could be beneficial for vector production and to further study their effect on vector assembly and budding.

8. References

AIUTI, A., BIASCO, L., SCARAMUZZA, S., FERRUA, F., CICALESE, M. P., BARICORDI, C., DIONISIO, F., CALABRIA, A., GIANNELLI, S., CASTIELLO, M. C., BOSTICARDO, M., EVANGELIO, C., ASSANELLI, A., CASIRAGHI, M., DI NUNZIO, S., CALLEGARO, L., BENATI, C., RIZZARDI, P., PELLIN, D., DI SERIO, C., SCHMIDT, M., VON KALLE, C., GARDNER, J., MEHTA, N., NEDUVA, V., DOW, D. J., GALY, A., MINIERO, R., FINOCCHI, A., METIN, A., BANERJEE, P. P., ORANGE, J. S., GALIMBERTI, S., VALSECCHI, M. G., BIFFI, A., MONTINI, E., VILLA, A., CICERI, F., RONCAROLO, M. G. & NALDINI, L. 2013. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science,* 341**,** 1233151.

AIUTI, A., CATTANEO, F., GALIMBERTI, S., BENNINGHOFF, U., CASSANI, B., CALLEGARO, L., SCARAMUZZA, S., ANDOLFI, G., MIROLO, M., BRIGIDA, I., TABUCCHI, A., CARLUCCI, F., EIBL, M., AKER, M., SLAVIN, S., AL-MOUSA, H., AL GHONAIUM, A., FERSTER, A., DUPPENTHALER, A., NOTARANGELO, L., WINTERGERST, U., BUCKLEY, R. H., BREGNI, M., MARKTEL, S., VALSECCHI, M. G., ROSSI, P., CICERI, F., MINIERO, R., BORDIGNON, C. & RONCAROLO, M. G. 2009. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *New England Journal of Medicine,* 360**,** 447-58.

AIUTI, A., SLAVIN, S., AKER, M., FICARA, F., DEOLA, S., MORTELLARO, A., MORECKI, S., ANDOLFI, G., TABUCCHI, A., CARLUCCI, F., MARINELLO, E., CATTANEO, F., VAI, S., SERVIDA, P., MINIERO, R., RONCAROLO, M. G. & BORDIGNON, C. 2002. Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science,* 296**,** 2410-3.

ARAVIN, A. A., LAGOS-QUINTANA, M., YALCIN, A., ZAVOLAN, M., MARKS, D., SNYDER, B., GAASTERLAND, T., MEYER, J. & TUSCHL, T. 2003. The small RNA profile during Drosophila melanogaster development. *Developmental Cell,* 5**,** 337-50.

ARHEL, N. J., SOUQUERE-BESSE, S., MUNIER, S., SOUQUE, P., GUADAGNINI, S., RUTHERFORD, S., PREVOST, M. C., ALLEN, T. D. & CHARNEAU, P. 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *The EMBO Journal,* 26**,** 3025-37.

ARUMUGAM, P. I., SCHOLES, J., PERELMAN, N., XIA, P., YEE, J. K. & MALIK, P. 2007. Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element. *Molecular Therapy,* 15**,** 1863-1871.

ASHKENAZI, A. & SHAI, Y. 2011. Insights into the mechanism of HIV-1 envelope induced membrane fusion as revealed by its inhibitory peptides. *Eur Biophys J.* 2011/01/25 ed.

BABST, M., ODORIZZI, G., ESTEPA, E. J. & EMR, S. D. 2000. Mammalian tumor susceptibility gene 101 (TSG101) and the yeast homologue, Vps23p, both function in late endosomal trafficking. *Traffic,* 1**,** 248-58.

BAEKELANDT, V., EGGERMONT, K., MICHIELS, M., NUTTIN, B. & DEBYSER, Z. 2003. Optimized lentiviral vector production and purification procedure prevents immune response after transduction of mouse brain. *Gene Therapy,* 10**,** 1933-1940.

BALASUBRAMANIAM, M. & FREED, E. O. 2011. New insights into HIV assembly and trafficking. *Physiology (Bethesda),* 26**,** 236-51.

BARR, J. N., WHELAN, S. P. J. & WERTZ, G. W. 2002. Transcriptional control of the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Biochimica Et Biophysica Acta-Gene Structure and Expression,* 1577**,** 337-353.

BAUBY, H., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE., INSERM, U., PARIS, FRANCE., LOPEZ-VERGÈS, S., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE., INSERM, U., PARIS, FRANCE., HOEFFEL, G., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE., INSERM, U., PARIS, FRANCE., DELCROIX-GENÊTE, D., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE., INSERM, U., PARIS,

FRANCE., JANVIER, K., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE., INSERM, U., PARIS, FRANCE., MAMMANO, F., INSERM U552, H. B. C. B., PARIS, FRANCE., HOSMALIN, A., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE., INSERM, U., PARIS, FRANCE., BERLIOZ-TORRENT, C., INSTITUT COCHIN, U. P. D., CNRS (UMR 8104), PARIS, FRANCE. & INSERM, U., PARIS, FRANCE. 2012. TIP47 is Required for the Production of Infectious HIV-1 Particles from Primary Macrophages. *Traffic,* 11**,** 455-467.

BAUM, C., HEGEWISCH-BECKER, S., ECKERT, H. G., STOCKING, C. & OSTERTAG, W. 1995. Novel retroviral vectors for efficient expression of the multidrug resistance (mdr-1) gene in early hematopoietic cells. *Journal of Virology,* 69**,** 7541-7.

BAUM, C., SCHAMBACH, A., BOHNE, J. & GALLA, M. 2006. Retrovirus vectors: toward the plentivirus? *Molecular Therapy,* 13**,** 1050-63.

BEAUSÉJOUR, TREMBLAY, M. J., BEAUSÉJOUR, Y. & TREMBLAY, M. J. 2004. Interaction between the Cytoplasmic Domain of ICAM-1 and Pr55Gag Leads to Acquisition of Host ICAM-1 by Human Immunodeficiency Virus Type 1. *Journal of Virology,* 78**,** 11916-11925.

BELL, A. J., FEGEN, D., WARD, M. & BANK, A. 2010. RD114 envelope proteins provide an effective and versatile approach to pseudotype lentiviral vectors. *Experimental Biology and Medicine (Maywood),* 235**,** 1269-1276.

BENABDELLAH, K., GUTIERREZ-GUERRERO, A., COBO, M., MUNOZ, P. & MARTIN, F. 2014. A Chimeric HS4-SAR Insulator (IS2) That Prevents Silencing and Enhances Expression of Lentiviral Vectors in Pluripotent Stem Cells. *PLoS One,* 9.

BERLIOZ-TORRENT, C., SHACKLETT, B. L., ERDTMANN, L., DELAMARRE, L., BOUCHAERT, I., SONIGO, P., DOKHELAR, M. C. & BENAROUS, R. 1999. Interactions of the cytoplasmic domains of human and simian retroviral transmembrane proteins with components of the clathrin adaptor complexes modulate intracellular and cell surface expression of envelope glycoproteins. *Journal of Virology,* 73**,** 1350-1361.

BERNSTEIN, E., CAUDY, A. A., HAMMOND, S. M. & HANNON, G. J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature,* 409**,** 363-6.

BESS, J. W., JR., GORELICK, R. J., BOSCHE, W. J., HENDERSON, L. E. & ARTHUR, L. O. 1997. Microvesicles are a source of contaminating cellular proteins found in purified HIV-1 preparations. *Virology,* 230**,** 134-144.

BIANCHI, M., HAKKIM, A., BRINKMANN, V., SILER, U., SEGER, R. A., ZYCHLINSKY, A. & REICHENBACH, J. 2009. Restoration of NET formation by gene therapy in CGD controls aspergillosis. *Blood,* 114**,** 2619-22.

BIANCHI, M., NIEMIEC, M. J., SILER, U., URBAN, C. F. & REICHENBACH, J. 2011. Restoration of anti-Aspergillus defense by neutrophil extracellular traps in human chronic granulomatous disease after gene therapy is calprotectin-dependent. *J Allergy Clin Immunol,* 127**,** 1243-52.e7.

BIFFI, A., MONTINI, E., LORIOLI, L., CESANI, M., FUMAGALLI, F., PLATI, T., BALDOLI, C., MARTINO, S., CALABRIA, A., CANALE, S., BENEDICENTI, F., VALLANTI, G., BIASCO, L., LEO, S., KABBARA, N., ZANETTI, G., RIZZO, W. B., MEHTA, N. A., CICALESE, M. P., CASIRAGHI, M., BOELENS, J. J., DEL CARRO, U., DOW, D. J., SCHMIDT, M., ASSANELLI, A., NEDUVA, V., DI SERIO, C., STUPKA, E., GARDNER, J., VON KALLE, C., BORDIGNON, C., CICERI, F., ROVELLI, A., RONCAROLO, M. G., AIUTI, A., SESSA, M. & NALDINI, L. 2013. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science,* 341**,** 1233158.

BINLEY, K., WIDDOWSON, P., LOADER, J., KELLEHER, M., IQBALL, S., FERRIGE, G., DE BELIN, J., CARLUCCI, M., ANGELL-MANNING, D., HURST, F., ELLIS, S., MISKIN, J., FERNANDES, A., WONG, P., ALLIKMETS, R., BERGSTROM, C., AABERG, T., YAN, J., KONG, J., GOURAS, P., PREFONTAINE, A., VEZINA, M., BUSSIERES, M., NAYLOR, S. & MITROPHANOUS, K. A. 2013. Transduction of Photoreceptors With Equine Infectious Anemia Virus Lentiviral Vectors: Safety and Biodistribution of StarGen for Stargardt Disease. *Investigative Ophthalmology & Visual Science,* 54**,** 4061-4071.

BINLEY, K., WIDDOWSON, P. S., KELLEHER, M., DE BELIN, J., LOADER, J., FERRIGE, G., CARLUCCI, M., ESAPA, M., CHIPCHASE, D., ANGELL-MANNING, D., ELLIS, S., MITROPHANOUS, K., MISKIN, J., BANTSEEV, V., NORK, T. M., MILLER, P. & NAYLOR, S. 2012. Safety and Biodistribution of an Equine Infectious Anemia Virus-Based Gene Therapy, RetinoStat((R)), for Age-Related Macular Degeneration. *Hum Gene Ther,* 23**,** 980-91.

BIRKENBACH, M., JOSEFSEN, K., YALAMANCHILI, R., LENOIR, G. & KIEFF, E. 1993. Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors. *Journal of Virology,* 67**,** 2209-2220.

BLACK, L. R. & AIKEN, C. 2010. TRIM5-Alpha Disrupts the Structure of Assembled HIV-1 Capsid Complexes In Vitro. *Journal of Virology,* 84**,** 6564-9.

BLOT, G., LOPEZ-VERGES, S., TREAND, C., KUBAT, N. J., DELCROIX-GENETE, D., EMILIANI, S., BENAROUS, R. & BERLIOZ-TORRENT, C. 2006. Luman, a new partner of HIV-1 TMgp41, interferes with tat-mediated transcription of the HIV-1 LTR. *Journal of Molecular Biology,* 364**,** 1034-1047.

BOCH, J., SCHOLZE, H., SCHORNACK, S., LANDGRAF, A., HAHN, S., KAY, S., LAHAYE, T., NICKSTADT, A. & BONAS, U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science,* 326**,** 1509-12.

BORDIGNON, C., NOTARANGELO, L. D., NOBILI, N., FERRARI, G., CASORATI, G., PANINA, P., MAZZOLARI, E., MAGGIONI, D., ROSSI, C., SERVIDA, P., UGAZIO, A. G. & MAVILIO, F. 1995. Gene therapy in peripheral blood lymphocytes and bone marrow for ADA-immunodeficient patients. *Science,* 270**,** 470-5.

BORGONOVO, B., COCUCCI, E., RACCHETTI, G., PODINI, P., BACHI, A. & MELDOLESI, J. 2002. Regulated exocytosis: a novel, widely expressed system. *Nature Cell Biology,* 4**,** 955-963.

BOSE, S., MATHUR, M., BATES, P., JOSHI, N. & BANERJEE, A. K. 2003. Requirement for cyclophilin A for the replication of vesicular stomatitis virus New Jersey serotype. *Journal of General Virology,* 84**,** 1687-99.

BOZTUG, K., SCHMIDT, M., SCHWARZER, A., BANERJEE, P. P., DÍEZ, I. A., DEWEY, R. A., BÖHM, M., NOWROUZI, A., BALL, C. R., GLIMM, H., NAUNDORF, S., KÜHLCKE, K., BLASCZYK, R., KONDRATENKO, I., MARÓDI, L., ORANGE, J. S., VON KALLE, C. & KLEIN, C. 2010. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *New England Journal of Medicine,* 363**,** 1918-27.

BRAATEN, D., FRANKE, E. K. & LUBAN, J. 1996. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *Journal of Virology,* 70**,** 3551-60.

BRASS, A. L., DYKXHOORN, D. M., BENITA, Y., YAN, N., ENGELMAN, A., XAVIER, R. J., LIEBERMAN, J. & ELLEDGE, S. J. 2008. Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. *Science*, 319, 921-6.

BRAUN, J.B., BOZTUG, K., PARUZYNSKI, A., WITZEL, M., SCHWARZER, A., ROTHE, M., MODLICH, U., BEIER, U., GÖHRING, G., STEINEMANN, D., FRONZA, R., BALL, C.R., HAEMMERLE, R., NAUNDORF, S., KÜHLCKE, K., ROSE, M., FRASER, C., MATHIAS, L., FERRARI, R., ABBOUD, M.R., AL-HERZ, W., KONDRATENKO, I., MARÓDI, L., GLIMM, H., SCHLEGELBERGER,B., SCHAMBACH, A., ALBERT, M.H., SCHMIDT, M., VON KALLE, C., KLEIN, C., 2014. Gene Therapy for Wiskott-Aldrich Syndrome - Long-Term Efficacy and Genotoxicity. Science Translational Medicine, 6, 227ra33

BRENNER, S., WHITING-THEOBALD, N. L., LINTON, G. F., HOLMES, K. L., ANDERSON-COHEN, M., KELLY, P. F., VANIN, E. F., PILON, A. M., BODINE, D. M., HORWITZ, M. E. & MALECH, H. L. 2003. Concentrated RD114-pseudotyped MFGS-gp91phox vector achieves high levels of functional correction of the chronic granulomatous disease oxidase defect in NOD/SCID/beta -microglobulin-/- repopulating mobilized human peripheral blood CD34+ cells. *Blood,* 102**,** 2789-97.

BRIGGS, J. A. & KRAUSSLICH, H. G. 2011. The molecular architecture of HIV. *Journal of Molecular Biology,* 410**,** 491-500.

BROUSSAU, S., JABBOUR, N., LACHAPELLE, G., DUROCHER, Y., TOM, R., TRANSFIGURACION, J., GILBERT, R. & MASSIE, B. 2008. Inducible Packaging Cells for Large-scale Production of Lentiviral Vectors in Serum-free Suspension Culture. *Molecular Therapy: The Journal of the American Society of Gene Therapy,* 16**,** 500-507.

BUCKLEY, R. H., SCHIFF, R. I., SCHIFF, S. E., MARKERT, M. L., WILLIAMS, L. W., HARVILLE, T. O., ROBERTS, J. L. & PUCK, J. M. 1997. Human severe combined immunodeficiency: genetic, phenotypic, and functional diversity in one hundred eight infants. *Journal of Pediatrics,* 130**,** 378-87.

BURNS, J. C., FRIEDMANN, T., DRIEVER, W., BURRASCANO, M. & YEE, J. K. 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 8033-8037.

BUSHMAN, F. D., MALANI, N., FERNANDES, J., D'ORSO, I., CAGNEY, G., DIAMOND, T. L., ZHOU, H., HAZUDA, D. J., ESPESETH, A. S., KONIG, R., BANDYOPADHYAY, S., IDEKER, T., GOFF, S. P., KROGAN, N. J., FRANKEL, A. D., YOUNG, J. A. & CHANDA, S. K. 2009. Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathogens,* 5**,** e1000437.

CAMUS, G., SEGURA-MORALES, C., MOLLE, D., LOPEZ-VERGES, S., BEGON-PESCIA, C., CAZEVIEILLE, C., SCHU, P., BERTRAND, E., BERLIOZ-TORRENT, C. & BASYUK, E. 2007. The clathrin adaptor complex AP-1 binds HIV-1 and MLV Gag and facilitates their budding. *Molelular Biolology of the Cell,* 18**,** 3193-3203.

CARLSON, L. A., BRIGGS, J. A., GLASS, B., RICHES, J. D., SIMON, M. N., JOHNSON, M. C., MULLER, B., GRUNEWALD, K. & KRAUSSLICH, H. G. 2008. Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis. *Cell Host Microbe,* 4**,** 592-599.

CARPENTIER, D. C., VEVIS, K., TRABALZA, A., GEORGIADIS, C., ELLISON, S. M., ASFAHANI, R. I. & MAZARAKIS, N. D. 2012. Enhanced pseudotyping efficiency of HIV-1 lentiviral vectors by a rabies/vesicular stomatitis virus chimeric envelope glycoprotein. *Gene Therapy,* 19**,** 761-74.

CARTIER, N., HACEIN-BEY-ABINA, S., BARTHOLOMAE, C. C., VERES, G., SCHMIDT, M., KUTSCHERA, I., VIDAUD, M., ABEL, U., DAL-CORTIVO, L., CACCAVELLI, L., MAHLAOUI, N., KIERMER, V., MITTELSTAEDT, D., BELLESME, C., LAHLOU, N., LEFRÈRE, F., BLANCHE, S., AUDIT, M., PAYEN, E., LEBOULCH, P., L'HOMME, B., BOUGNÈRES, P., VON KALLE, C., FISCHER, A., CAVAZZANA-CALVO, M. & AUBOURG, P. 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science,* 326**,** 818-823.

CAVAZZANA-CALVO, M., HACEIN-BEY, S., DE SAINT BASILE, G., GROSS, F., YVON, E., NUSBAUM, P., SELZ, F., HUE, C., CERTAIN, S., CASANOVA, J. L., BOUSSO, P., DEIST, F. L. & FISCHER, A. 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science,* 288**,** 669-72.

CAVAZZANA-CALVO, M., PAYEN, E., NEGRE, O., WANG, G., HEHIR, K., FUSIL, F., DOWN, J., DENARO, M., BRADY, T., WESTERMAN, K., CAVALLESCO, R., GILLET-LEGRAND, B., CACCAVELLI, L., SGARRA, R., MAOUCHE-CHRÉTIEN, L., BERNAUDIN, F., GIROT, R., DORAZIO, R., MULDER, G.-J., POLACK, A., BANK, A., SOULIER, J., LARGHERO, J., KABBARA, N., DALLE, B., GOURMEL, B., SOCIE, G., CHRÉTIEN, S., CARTIER, N., AUBOURG, P., FISCHER, A., CORNETTA, K., GALACTEROS, F., BEUZARD, Y., GLUCKMAN, E., BUSHMAN, F., HACEIN-BEY-ABINA, S. & LEBOULCH, P. 2010. Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. *Nature,* 467**,** 318-322.

CHATEL-CHAIX, L., ABRAHAMYAN, L., FRECHINA, C., MOULAND, A. J. & DESGROSEILLERS, L. 2007. The host protein Staufen1 participates in human immunodeficiency virus type 1

assembly in live cells by influencing pr55Gag multimerization. *Journal of Virology,* 81**,** 6216-30.

CHECKLEY, M. A., LUTTGE, B. G. & FREED, E. O. 2011. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *Journal of Molecular Biology,* 410**,** 582-608.

CHERTOVA, E., CHERTOV, O., COREN, L. V., ROSER, J. D., TRUBEY, C. M., BESS, J. W., JR., SOWDER, R. C., 2ND, BARSOV, E., HOOD, B. L., FISHER, R. J., NAGASHIMA, K., CONRADS, T. P., VEENSTRA, T. D., LIFSON, J. D. & OTT, D. E. 2006. Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. *Journal of Virology,* 80**,** 9039-9052.

CHINEN, J., DAVIS, J., DE RAVIN, S. S., HAY, B. N., HSU, A. P., LINTON, G. F., NAUMANN, N., NOMICOS, E. Y. H., SILVIN, C., ULRICK, J., WHITING-THEOBALD, N. L., MALECH, H. L. & PUCK, J. M. 2007. Gene therapy improves immune function in preadolescents with X-linked severe combined immunodeficiency. *Blood,* 110**,** 67-73.

CHUNG, C. S., CHEN, C. H., HO, M. Y., HUANG, C. Y., LIAO, C. L. & CHANG, W. 2006. Vaccinia virus proteome: identification of proteins in vaccinia virus intracellular mature virion particles. *Journal of Virology,* 80**,** 2127-40.

CHUNG, J. H., BELL, A. C. & FELSENFELD, G. 1997. Characterization of the chicken beta-globin insulator. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 575-580.

CIMARELLI, A. & LUBAN, J. 1999. Translation elongation factor 1-alpha interacts specifically with the human immunodeficiency virus type 1 Gag polyprotein. *Journal of Virology,* 73**,** 5388-401.

CLAVEL, F. & CHARNEAU, P. 1994. Fusion from without directed by human immunodeficiency virus particles. *Journal of Virology,* 68**,** 1179-85.

CLINICALTRIALS.GOV. 2013. *Gene Therapy for X-linked Chronic Granulomatous Disease (X-CGD) - Full Text View - ClinicalTrials.gov* [Online]. Available: http://clinicaltrials.gov/show/NCT01855685.

COFFIN, J. M., H HUGHES, S. & VARMUS, H. E. 1997. Retroviral "Lifestyles": Simple versus Complex. *Retroviruses.* http://www.ncbi.nlm.nih.gov/books/NBK19363/: Cold Spring Harbor Laboratory Press.

COIL, D. A. & MILLER, A. D. 2004. Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. *Journal of Virology,* 78**,** 10920-6.

COLEMAN, J. E., HUENTELMAN, M. J., KASPAROV, S., METCALFE, B. L., PATON, J. F., KATOVICH, M. J., SEMPLE-ROWLAND, S. L. & RAIZADA, M. K. 2003. Efficient large-scale production and concentration of HIV-1-based lentiviral vectors for use in vivo. *Physiological Genomics,* 12**,** 221-8.

CORNETTA, K., POLLOK, K. E. & MILLER, A. D. 2008. Transduction of Primary Hematopoietic Cells by Retroviral Vectors. *Cold Spring Harbor Laboratory Press*, 34, e120.

COSSET, F. L., TAKEUCHI, Y., BATTINI, J. L., WEISS, R. A. & COLLINS, M. K. 1995. High-titer packaging cells producing recombinant retroviruses resistant to human serum. *Journal of Virology,* 69**,** 7430-6.

CRUZ, P. E., GONÇALVES, D., ALMEIDA, J., MOREIRA, J. L. & CARRONDO, M. J. T. 2000. Modeling retrovirus production for gene therapy. 2. Integrated optimization of bioreaction and downstream processing. *Biotechnology Progress,* 16**,** 350-357.

CURETON, D. K., MASSOL, R. H., SAFFARIAN, S., KIRCHHAUSEN, T. L. & WHELAN, S. P. 2009. Vesicular stomatitis virus enters cells through vesicles incompletely coated with clathrin that depend upon actin for internalization. *PLoS Pathog,* 5**,** e1000394.

DAS, S. C., NAYAK, D., ZHOU, Y. & PATTNAIK, A. K. 2006. Visualization of intracellular transport of vesicular stomatitis virus nucleocapsids in living cells. *Journal of Virology,* 80**,** 6368-77.

DAVIS, D. A., NEWCOMB, F. M., STARKE, D. W., OTT, D. E., MIEYAL, J. J. & YARCHOAN, R. 1997. Thioltransferase (glutaredoxin) is detected-within HIV-1 and can regulate the activity

of glutathionylated HIV-1 protease in vitro. *Journal of Biological Chemistry,* 272**,** 25935-25940.

DENARD, J., RUNDWASSER, S., LAROUDIE, N., GONNET, F., NALDINI, L., RADRIZZANI, M., GALY, A., MERTEN, O.-W., DANOS, O. & SVINARTCHOUK, F. 2009. Quantitative proteomic analysis of lentiviral vectors using 2-DE. *PROTEOMICS,* 9**,** 3666-3676.

DENEKA, M., PELCHEN-MATTHEWS, A., BYLAND, R., RUIZ-MATEOS, E. & MARSH, M. 2007. In macrophages, HIV-1 assembles into an intracellular plasma membrane domain containing the tetraspanins CD81, CD9, and CD53. *Journal of Cell Biology,* 177**,** 329-41.

DONG, X., LI, H., DERDOWSKI, A., DING, L., BURNETT, A., CHEN, X., PETERS, T. R., DERMODY, T. S., WOODRUFF, E., WANG, J. J. & SPEARMAN, P. 2005. AP-3 directs the intracellular trafficking of HIV-1 Gag and plays a key role in particle assembly. *Cell,* 120**,** 663-74.

DOOHER, J. E., SCHNEIDER, B. L., REED, J. C. & LINGAPPA, J. R. 2007. Host ABCE1 is at plasma membrane HIV assembly sites and its dissociation from Gag is linked to subsequent events of virus production. *Traffic,* 8**,** 195-211.

DULL, T., ZUFFEREY, R., KELLY, M., MANDEL, R. J., NGUYEN, M., TRONO, D. & NALDINI, L. 1998. A Third-Generation Lentivirus Vector with a Conditional Packaging System. *Journal of Virology,* 72**,** 8463-8471.

EDELING, M. A., SMITH, C. & OWEN, D. 2006. Life of a clathrin coat: insights from clathrin and AP structures. *Nature Reviews Molecular Cell Biology,* 7**,** 32-44.

ELLIS, J. 2005. Silencing and variegation of gammaretrovirus and lentivirus vectors. *Human Gene Therapy,* 16**,** 1241-1246.

EMERSON, V., HOLTKOTTE, D., PFEIFFER, T., WANG, I. H., SCHNOLZER, M., KEMPF, T. & BOSCH, V. 2010. Identification of the Cellular Prohibitin 1/Prohibitin 2 Heterodimer as an Interaction Partner of the C-Terminal Cytoplasmic Domain of the HIV-1 Glycoprotein. *Journal of Virology,* 84**,** 1355-1365.

ENG, K. J., MCCORMARCK, L. A. & YATES, J. R. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry*, 5, 976–989.

EVANS, D. T., TILLMAN, K. C. & DESROSIERS, R. C. 2002. Envelope glycoprotein cytoplasmic domains from diverse lentiviruses interact with the prenylated Rab acceptor. *Journal of Virology,* 76**,** 327-337.

FARSON, D., WITT, R., MCGUINNESS, R., DULL, T., KELLY, M., SONG, J. P., RADEKE, R., BUKOVSKY, A., CONSIGLIO, A. & NALDINI, L. 2001. A new-generation stable inducible packaging cell line for lentiviral vectors. *Human Gene Therapy,* 12**,** 981-997.

FASSATI, A., GORLICH, D., HARRISON, I., ZAYTSEVA, L. & MINGOT, J. M. 2003. Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. *The EMBO Journal,* 22**,** 3675-85.

FENG, Y., BRODER, C. C., KENNEDY, P. E. & BERGER, E. A. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science,* 272**,** 872-7.

FENG, Y.-X., CAMPBELL, S., HARVIN, D., EHRESMANN, B., EHRESMANN, C., REIN, A., FENG, Y.-X., CAMPBELL, S., HARVIN, D., EHRESMANN, B., EHRESMANN, C. & REIN, A. 1999. The Human Immunodeficiency Virus Type 1 Gag Polyprotein Has Nucleic Acid Chaperone Activity: Possible Role in Dimerization of Genomic RNA and Placement of tRNA on the Primer Binding Site. *Journal of Virology,* 73**,** 4251-4256.

FERRUA, F., BRIGIDA, I. & AIUTI, A. 2010. Update on gene therapy for adenosine deaminase-deficient severe combined immunodeficiency. *Current Opinion in Allergy and Clinical Immunology,* 10**,** 551-6.

FINKELSHTEIN, D., NOVICK, A.D., BARAK, S. AND RUBINSTEIN, M., 2013. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. Proceedings of the National Academy of Sciences of the United States of America, 110(18): 7306–7311.

FIRE, A., XU, S., MONTGOMERY, M. K., KOSTAS, S. A., DRIVER, S. E. & MELLO, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature,* 391**,** 806-811.

FISCHER, A. 2000. Severe combined immunodeficiencies (SCID). *Clinical and Experimental Immunology,* 122**,** 143-9.

FOLLENZI, A., AILLES, L. E., BAKOVIC, S., GEUNA, M. & NALDINI, L. 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nature Genetics,* 25**,** 217-22.

FRANKE, E. K., YUAN, H. E. & LUBAN, J. 1994. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature,* 372**,** 359-62.

FRECHA, C., COSTA, C., LEVY, C., NEGRE, D., RUSSELL, S. J., MAISNER, A., SALLES, G., PENG, K.-W., COSSET, F.-L. & VERHOEYEN, E. 2009. Efficient and stable transduction of resting B lymphocytes and primary chronic lymphocyte leukemia cells using measles virus gp displaying lentiviral vectors. *Blood,* 114**,** 3173-3180.

FRECHA, C., COSTA, C., NEGRE, D., GAUTHIER, E., RUSSELL, S. J., COSSET, F. L. & VERHOEYEN, E. 2008. Stable transduction of quiescent T cells without induction of cycle progression by a novel lentiviral vector pseudotyped with measles virus glycoproteins. *Blood,* 112**,** 4843-52.

FUKUMA, A., ABE, M., URATA, S., YOSHIKAWA, R., MORIKAWA, Y., MIYAZAWA, T. & YASUDA, J. 2011. Viral and cellular requirements for the budding of feline endogenous retrovirus RD-114. *Virology Journal,* 8**,** 540.

GAJ, T., GUO, J., KATO, Y., SIRK, S. J. & BARBAS, C. F., 3RD 2012. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nature Methods,* 9**,** 805-7.

GARRUS, J. E., VON SCHWEDLER, U. K., PORNILLOS, O. W., MORHAM, S. G., ZAVITZ, K. H., WANG, H. E., WETTSTEIN, D. A., STRAY, K. M., COTE, M., RICH, R. L., MYSZKA, D. G. & SUNDQUIST, W. I. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell,* 107**,** 55-65.

GASPAR, H. B. 2012. Gene therapy for ADA-SCID: defining the factors for successful outcome. *Blood,* 120**,** 3628-9.

GASPAR, H. B., BJORKEGREN, E., PARSLEY, K., GILMOUR, K. C., KING, D., SINCLAIR, J., ZHANG, F., GIANNAKOPOULOS, A., ADAMS, S., FAIRBANKS, L. D., GASPAR, J., HENDERSON, L., XU-BAYFORD, J. H., DAVIES, E. G., VEYS, P. A., KINNON, C. & THRASHER, A. J. 2006. Successful Reconstitution of Immunity in ADA-SCID by Stem Cell Gene Therapy Following Cessation of PEG-ADA and Use of Mild Preconditioning. *Molecular Therapy,* 14**,** 505-513.

GASPAR, H. B., COORAY, S., GILMOUR, K., PARSLEY, K. L., ZHANG, F., ADAMS, S., BJORKEGREN, E., BAYFORD, J., BROWN, L., DAVIES, E. G., VEYS, P., FAIRBANKS, L., BORDON, V., PETROPOULOU, T., KINNON, C. & THRASHER, A. J. 2011. Hematopoietic Stem Cell Gene Therapy for Adenosine Deaminase–Deficient Severe Combined Immunodeficiency Leads to Long-Term Immunological Recovery and Metabolic Correction. *Science Translational Medicine*, 3, 97ra80.

GASPAR, H. B., PARSLEY, K. L., HOWE, S., KING, D., GILMOUR, K. C., SINCLAIR, J., BROUNS, G., SCHMIDT, M., VON KALLE, C., BARINGTON, T., JAKOBSEN, M. A., CHRISTENSEN, H. O., AL GHONAIUM, A., WHITE, H. N., SMITH, J. L., LEVINSKY, R. J., ALI, R. R., KINNON, C. & THRASHER, A. J. 2004. Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet,* 364**,** 2181-7.

GENETHON. 2013. *Genethon receives approval from the UK medicine regulatory agency to start a new clinical trial for an inherited immune deficiency: chronic granulomatous disease.* [Online]. Available: http://www.genethon.fr/en/2013/02/05/genethon-receives-approval-from-the-uk-medicine-regulatory-agency-to-start-a-new-clinical-trial-for-an-inherited-immune-deficiency-chronic-granulomatous-disease/ [Accessed 22 January 2014].

GERAERTS, M., MICHIELS, M., BAEKELANDT, V., DEBYSER, Z. & GIJSBERS, R. 2005. Upscaling of lentiviral vector production by tangential flow filtration. *The Journal of Gene Medicine,* 7**,** 1299-1310.

GHANI, K., GARNIER, A., COELHO, H., TRANSFIGURACION, J., TRUDEL, P. & KAMEN, A. 2006. Retroviral vector production using suspension-adapted 293GPG cells in a 3L acoustic filter-based perfusion bioreactor. *Biotechnology and Bioengineering,* 95**,** 653-60.

GHEYSEN, D. 1989. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell,* 59**,** 103-112.

GLUSCHANKOF, P., MONDOR, I., GELDERBLOM, H. R. & SATTENTAU, Q. J. 1997. Cell Membrane Vesicles Are a Major Contaminant of Gradient-Enriched Human Immunodeficiency Virus Type-1 Preparations. *Virology,* 230**,** 125-133.

GOJOBORI, T., MORIYAMA, E.N., INA, Y., IKEO, K., MIURA, T., TSUJIMOTO, H., HAYAMI, M. AND YOKOYAMA, S., 1990. Evolutionary origin of human and simian immunodeficiency viruses. Proceedings of the National Academy of Sciences of the United States of America, 87(11): 4108–4111.

Gousset, K., Ablan, S.D., Coren, L.V., Ono, A., Soheilian, F., Nagashima, K., Ott, D.E., Freed, E.O., 2008. Real-time visualization of HIV-1 GAG trafficking in infected macrophages. PLoS Pathogen, 4, 3, e1000015.

GRAHAM, D. R. M., CHERTOVA, E., HILBURN, J. M., ARTHUR, L. O. & HILDRETH, J. E. K. 2003. Cholesterol Depletion of Human Immunodeficiency Virus Type 1 and Simian Immunodeficiency Virus with ?-Cyclodextrin Inactivates and Permeabilizes the Virions: Evidence for Virion-Associated Lipid Rafts. *Journal of Virology,* 77**,** 8237-48.

GRAHAM, F. L., SMILEY, J., RUSSELL, W. C. & NAIRN, R. 1977. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology,* 36**,** 59-72.

GRUPP, S. A., KALOS, M., BARRETT, D., APLENC, R., PORTER, D. L., RHEINGOLD, S. R., TEACHEY, D. T., CHEW, A., HAUCK, B., WRIGHT, J. F., MILONE, M. C., LEVINE, B. L. & JUNE, C. H. 2013. Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. *New England Journal of Medicine,* 368**,** 1509-18.

GURER, C., CIMARELLI, A. & LUBAN, J. 2002. Specific incorporation of heat shock protein 70 family members into primate lentiviral virions. *Journal of Virology,* 76**,** 4666-4670.

HACEIN-BEY, H., CAVAZZANA-CALVO, M., LE DEIST, F., DAUTRY-VARSAT, A., HIVROZ, C., RIVIERE, I., DANOS, O., HEARD, J. M., SUGAMURA, K., FISCHER, A. & DE SAINT BASILE, G. 1996. gamma-c gene transfer into SCID X1 patients' B-cell lines restores normal high-affinity interleukin-2 receptor expression and function. *Blood,* 87**,** 3108-16.

HACEIN-BEY-ABINA, S., GARRIGUE, A., WANG, G. P., SOULIER, J., LIM, A., MORILLON, E., CLAPPIER, E., CACCAVELLI, L., DELABESSE, E., BELDJORD, K., ASNAFI, V., MACINTYRE, E., DAL CORTIVO, L., RADFORD, I., BROUSSE, N., SIGAUX, F., MOSHOUS, D., HAUER, J., BORKHARDT, A., BELOHRADSKY, B. H., WINTERGERST, U., VELEZ, M. C., LEIVA, L., SORENSEN, R., WULFFRAAT, N., BLANCHE, S., BUSHMAN, F. D., FISCHER, A. & CAVAZZANA-CALVO, M. 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *The Journal of Clinical Investigation,* 118**,** 3132-42.

HACEIN-BEY-ABINA, S., HAUER, J., LIM, A., PICARD, C., WANG, G. P., BERRY, C. C., MARTINACHE, C., RIEUX-LAUCAT, F., LATOUR, S., BELOHRADSKY, B. H., LEIVA, L., SORENSEN, R., DEBRE, M., CASANOVA, J. L., BLANCHE, S., DURANDY, A., BUSHMAN, F. D., FISCHER, A. & CAVAZZANA-CALVO, M. 2010. Efficacy of gene therapy for X-linked severe combined immunodeficiency. *New England Journal of Medicine,* 363**,** 355-64.

HACEIN-BEY-ABINA, S., LE DEIST, F., CARLIER, F., BOUNEAUD, C., HUE, C., DE VILLARTAY, J. P., THRASHER, A. J., WULFFRAAT, N., SORENSEN, R., DUPUIS-GIROD, S., FISCHER, A., DAVIES, E. G., KUIS, W., LEIVA, L. & CAVAZZANA-CALVO, M. 2002. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *New England Journal of Medicine,* 346**,** 1185-93.

HALLENBERGER, S., BOSCH, V., ANGLIKER, H., SHAW, E., KLENK, H.-D. & GARTEN, W. 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gpl60. *Nature*, 360**,** 358-361.

HAMILTON, A. J. & BAULCOMBE, D. C. 1999. A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. *Science*, 286, 950-2

HAMMOND, S. M., BERNSTEIN, E., BEACH, D. & HANNON, G. J. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature,* 404**,** 293-6.

HANTZOPOULOS, P. A., SULLENGER, B. A., UNGERS, G. & GILBOA, E. 1989. Improved gene expression upon transfer of the adenosine deaminase minigene outside the transcriptional unit of a retroviral vector. *Proceedings of the National Academy of Sciences of the United States of America,* 86**,** 3519-23.

HARRIS, R. S., BISHOP, K. N., SHEEHY, A. M., CRAIG, H. M., PETERSEN-MAHRT, S. K., WATT, I. N., NEUBERGER, M. S. & MALIM, M. H. 2003. DNA deamination mediates innate immunity to retroviral infection. *Cell,* 113**,** 803-9.

HASHIMOTO, T., GAMOU, S., SHIMIZU, N., KITAJIMA, Y. & NISHIKAWA, T. 1995. Regulation of translocation of the desmoyokin/AHNAK protein to the plasma membrane in keratinocytes by protein kinase C. *Experimental Cell Research,* 217**,** 258-66.

HELSETH, E., KOWALSKI, M., GABUZDA, D., OLSHEVSKY, U., HASELTINE, W. & SODROSKI, J. 1990. RAPID COMPLEMENTATION ASSAYS MEASURING REPLICATIVE POTENTIAL OF HUMAN-IMMUNODEFICIENCY-VIRUS TYPE-1 ENVELOPE GLYCOPROTEIN MUTANTS. *Journal of Virology,* 64**,** 2416-2420.

HIEDA, Y. & TSUKITA, S. 1989. A new high molecular mass protein showing unique localization in desmosomal plaque. *Journal of Cell Biology,* 109**,** 1511-8.

HILDINGER, M., ABEL, K. L., OSTERTAG, W. & BAUM, C. 1999. Design of 5' untranslated sequences in retroviral vectors developed for medical use. *Journal of Virology,* 73**,** 4083-9.

HOHAUS, A., PERSON, V., BEHLKE, J., SCHAPER, J., MORANO, I. & HAASE, H. 2002. The carboxyl-terminal region of ahnak provides a link between cardiac L-type Ca2+. The FASEB Journal*,* 16**,** 1205-16.

HOWE, S. J., MANSOUR, M. R., SCHWARZWAELDER, K., BARTHOLOMAE, C., HUBANK, M., KEMPSKI, H., BRUGMAN, M. H., PIKE-OVERZET, K., CHATTERS, S. J., DE RIDDER, D., GILMOUR, K. C., ADAMS, S., THORNHILL, S. I., PARSLEY, K. L., STAAL, F. J., GALE, R. E., LINCH, D. C., BAYFORD, J., BROWN, L., QUAYE, M., KINNON, C., ANCLIFF, P., WEBB, D. K., SCHMIDT, M., VON KALLE, C., GASPAR, H. B. & THRASHER, A. J. 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *The Journal of Clinical Investigations,* 118**,** 3143-50.

HUANG, M., ORENSTEIN, J. M., MARTIN, M. A. & FREED, E. O. 1995. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *Journal of Virology,* 69**,** 6810-8.

HULME, A. E., PEREZ, O. & HOPE, T. J. 2011. Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 9975-9980.

IKEDA, Y., TAKEUCHI, Y., MARTIN, F., COSSET, F. L., MITROPHANOUS, K. & COLLINS, M. 2003. Continuous high-titer HIV-1 vector production. *Nature Biotechnology,* 21**,** 569-72.

JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity.

JOHANNSDOTTIR, H. K., MANCINI, R., KARTENBECK, J., AMATO, L. & HELENIUS, A. 2009. Host cell factors and functions involved in vesicular stomatitis virus entry. *Journal of Virology,* 83**,** 440-53.

JOHANNSEN, E., LUFTIG, M., CHASE, M. R., WEICKSEL, S., CAHIR-MCFARLAND, E., ILLANES, D., SARRACINO, D. & KIEFF, E. 2004. Proteins of purified Epstein-Barr virus. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 16286-91.

JOLLY, C. 2010. T Cell Polarization at the Virological Synapse. *Viruses,* 2**,** 1261-1278.

JOLLY, C., MITAR, I. & SATTENTAU, Q. J. 2007. Requirement for an intact T-cell actin and tubulin cytoskeleton for efficient assembly and spread of human immunodeficiency virus type 1. *Journal of Virology,* 81**,** 5547-5560.

JORDAN, A., DEFECHEREUX, P. & VERDIN, E. 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *The EMBO Journal,* 20**,** 1726-1738.

JORDAN, M. & WURM, F. 2004. Transfection of adherent and suspended cells by calcium phosphate. *Methods,* 33**,** 136-43.

JOUVE, M., SOL-FOULON, N., WATSON, S., SCHWARTZ, O. & BENAROCH, P. 2007. HIV-1 Buds and Accumulates in "Nonacidic" Endosomes of Macrophages. *Cell Host & Microbe,* 2**,** 85-95.

JOUVENET, N., AARON DIAMOND AIDS RESEARCH CENTER, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, LABORATORY OF RETROVIROLOGY, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, NEIL, S. J. D., AARON DIAMOND AIDS RESEARCH CENTER, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, LABORATORY OF RETROVIROLOGY, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, BESS, C., LABORATORY OF CELLULAR BIOPHYSICS, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, JOHNSON, M. C., DEPARTMENT OF MOLECULAR MICROBIOLOGY AND IMMUNOLOGY, U. O. M.-C., COLUMBIA, MISSOURI, UNITED STATES OF AMERICA, VIRGEN, C. A., AARON DIAMOND AIDS RESEARCH CENTER, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, LABORATORY OF RETROVIROLOGY, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, SIMON, S. M., LABORATORY OF CELLULAR BIOPHYSICS, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA, BIENIASZ, P. D., AARON DIAMOND AIDS RESEARCH CENTER, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA & LABORATORY OF RETROVIROLOGY, T. R. U., NEW YORK, NEW YORK, UNITED STATES OF AMERICA 2006. Plasma Membrane Is the Site of Productive HIV-1 Particle Assembly. *PLoS Biology,* 4.

KANG, E. M., CHOI, U., THEOBALD, N., LINTON, G., LONG PRIEL, D. A., KUHNS, D. & MALECH, H. L. 2010. Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. *Blood,* 115**,** 783-91.

KANG, H. J., BARTHOLOMAE, C. C., PARUZYNSKI, A., ARENS, A., KIM, S., YU, S. S., HONG, Y., JOO, C. W., YOON, N. K., RHIM, J. W., KIM, J. G., VON KALLE, C., SCHMIDT, M. & AHN, H. S. 2011. Retroviral gene therapy for X-linked chronic granulomatous disease: results from phase I/II trial. *Molecular Therapy,* 19**,** 2092-101.

KARN, J. & STOLTZFUS, C. M. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor Perspectives in Medicine,* 2**,** a006916.

KATOH, K., SHIBATA, H., SUZUKI, H., NARA, A., ISHIDOH, K., KOMINAMI, E., YOSHIMORI, T. & MAKI, M. 2003. The ALG-2-interacting protein Alix associates with CHMP4b, a human homologue of yeast Snf7 that is involved in multivesicular body sorting. *Journal of Biological Chemistry,* 278**,** 39104-13.

KEATING, K., ZHAO, Y., DOLMAN, C. & THORPE, R. 2008. Characterization of complete particles (VSV-G/SIN-GFP) and empty particles (VSV-G/EMPTY) in human immunodeficiency virus type 1-based lentiviral products for gene therapy: potential applications for improvement of product quality and safety. *Human Gene Therapy,* 19**,** 475-486.

KIM, J. T., KIM, E. M., LEE, K. H., CHOI, J. E., JHUN, B. H. & KIM, J. W. 2002. Leucine zipper domain of HIV-1 gp41 interacted specifically with alpha-catenin. *Biochemical and Biophysical Research Communications,* 291**,** 1239-1244.

KIM, V. N., MITROPHANOUS, K., KINGSMAN, S. M. & KINGSMAN, A. J. 1998. Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1. *Journal of Virology,* 72**,** 811-816.

KIMPTON, J. & EMERMAN, M. 1992. Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. *Journal of Virology*, 66, 2232-39

KLAGES, N., ZUFFEREY, R. & TRONO, D. 2000. A stable system for the high-titer production of multiply attenuated lentiviral vectors. *Molecular Therapy: The Journal of the American Society of Gene Therapy,* 2**,** 170-176.

KNIGHT, S., ZHANG, F., MUELLER-KULLER, U., BOKHOVEN, M., GUPTA, A., BROUGHTON, T., SHA, S., ANTONIOU, M. N., BRENDEL, C., GREZ, M., THRASHER, A. J., COLLINS, M. & TAKEUCHI, Y. 2012. Safer, silencing-resistant lentiviral vectors: optimization of the ubiquitous chromatin-opening element through elimination of aberrant splicing. *Journal of Virology,* 86**,** 9088-95.

KOK, K. H., LEI, T. & JIN, D. Y. 2009. siRNA and shRNA screens advance key understanding of host factors required for HIV-1 replication. *Retrovirology,* 6**,** 78.

KOSAKA, Y., KOBAYASHI, N., FUKAZAWA, T., TOTSUGAWA, T., MARUYAMA, M., YONG, C., ARATA, T., IKEDA, H., KOBAYASHI, K., UEDA, T., KURABAYASHI, Y. & TANAKA, N. 2004. Lentivirus-based gene delivery in mouse embryonic stem cells. *Artificial Organs,* 28**,** 271-7.

KOTSOPOULOU, E., KIM, V. N., KINGSMAN, A. J., KINGSMAN, S. M. & MITROPHANOUS, K. A. 2000. A Rev-Independent Human Immunodeficiency Virus Type 1 (HIV-1)-Based Vector That Exploits a Codon-Optimized HIV-1 gag-pol Gene. *Journal of Virology,* 74**,** 4839-4852.

KRISHNAN, L. & ENGELMAN, A. 2012. Retroviral Integrase Proteins and HIV-1 DNA Integration. *Journal of Biological Chemistry,* 287**,** 40858-40866.

KUMAR, M., KELLER, B., MAKALOU, N. & SUTTON, R. E. 2001. Systematic determination of the packaging limit of lentiviral vectors. *Human Gene Therapy,* 12**,** 1893-905.

KUTAY, U., IZAURRALDE, E., BISCHOFF, F. R., MATTAJ, I. W. & GORLICH, D. 1997. Dominant-negative mutants of importin-beta block multiple pathways of import and export through the nuclear pore complex. *The EMBO Journal,* 16**,** 1153-1163.

KUTNER, R. H., ZHANG, X.-Y. & REISER, J. 2009. Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. *Nature Protocols,* 4**,** 495-505.

KÖNIG, R., ZHOU, Y., ELLEDER, D., DIAMOND, T. L., BONAMY, G. M., IRELAN, J. T., CHIANG, C. Y., TU, B. P., DE JESUS, P. D., LILLEY, C. E., SEIDEL, S., OPALUCH, A. M., CALDWELL, J. S., WEITZMAN, M. D., KUHEN, K. L., BANDYOPADHYAY, S., IDEKER, T., ORTH, A. P., MIRAGLIA, L. J., BUSHMAN, F. D., YOUNG, J. A. & CHANDA, S. K. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell,* 135**,** 49-60.

LANGEN, H., TAKACS, B., EVERS, S., PETER, B., HANS-WERNER, L., WIPF, B., CHRISTOPHER, G. & FOUNTOULAKIS, M. 2000. Two- dimensional map of the proteome of Haemophilus influenzae. *Electrophoresis,* 21**,** 411-429.

LARRIBA, S., SUMOY, L., RAMOS, M. D., GIMENEZ, J., ESTIVILL, X., CASALS, T. & NUNES, V. 2001. ATB(0)/SLC1A5 gene. Fine localisation and exclusion of association with the intestinal phenotype of cystic fibrosis. *European Journal of Human Genetics,* 9**,** 860-6.

LEAVITT, A. D., ROBLES, G., ALESANDRO, N. & VARMUS, H. E. 1996. Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. *Journal of Virology,* 70**,** 721-728.

LEBOULCH, P. 2013. Gene therapy: primed for take-off. *Nature, 500, 280-2*

LEE, C. L., CHOU, M., DAI, B., XIAO, L. & WANG, P. 2012. Construction of stable producer cells to make high-titer lentiviral vectors for dendritic cell-based vaccination. *Biotechnology and Bioengineering,* 109**,** 1551-60.

LEE, I. H., LIM, H. J., YOON, S., SEONG, J. K., BAE, D. S., RHEE, S. G. & BAE, Y. S. 2008. Ahnak protein activates protein kinase C (PKC) through dissociation of the PKC-protein phosphatase 2A complex. *Journal of Biological Chemistry,* 283**,** 6312-20.

LEE, R. C., FEINBAUM, R. L. & AMBROS, V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* 75**,** 843-54.

LEE, Y., AHN, C., HAN, J., CHOI, H., KIM, J., YIM, J., LEE, J., PROVOST, P., RADMARK, O., KIM, S. & KIM, V. N. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature,* 425**,** 415-9.

LEE, Y., JEON, K., LEE, J. T., KIM, S. & KIM, V. N. 2002. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal,* 21**,** 4663-70.

LEFKOWITZ, E. J., BAO, Y., BRISTER, J. R., DEMPSEY, D., HENDRICKSON, R. C., KUIKEN, C., LE MERCIER, P., SCHEUERMANN, R., SETO, D. & UPTON, C. 2013. *International Committee on Taxonomy of Viruses* [Online]. http://www.ictvonline.org/virusTaxonomy.asp: International Committee on Taxonomy of Viruses (ICTV). [Accessed 21 January 2014].

LEVINE, B. L., HUMEAU, L. M., BOYER, J., MACGREGOR, R. R., REBELLO, T., LU, X., BINDER, G. K., SLEPUSHKIN, V., LEMIALE, F., MASCOLA, J. R., BUSHMAN, F. D., DROPULIC, B. & JUNE, C. H. 2006. Gene transfer in humans using a conditionally replicating lentiviral vector. *Proceedings of the National Academy of Sciences of the United States of America,* 103, 17372-7.

LEWIS, P. F. & EMERMAN, M. 1994. PASSAGE THROUGH MITOSIS IS REQUIRED FOR ONCORETROVIRUSES BUT NOT FOR THE HUMAN-IMMUNODEFICIENCY-VIRUS. *Journal of Virology,* 68**,** 510-516.

LINGWOOD, D. & SIMONS, K. 2010. Lipid rafts as a membrane-organizing principle. *Science,* 327**,** 46-50.

LINK, A. J., ENG, J., SCHIELTZ, D. M., CARMACK, E., J, M. G., MORRIS, D. R., GARVIK, B. M. & YATES JOHN R, I. 1999. Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology,* 17**,** 676-682.

LOMBARDO, A., GENOVESE, P., BEAUSEJOUR, C. M., COLLEONI, S., LEE, Y. L., KIM, K. A., ANDO, D., URNOV, F. D., GALLI, C., GREGORY, P. D., HOLMES, M. C. & NALDINI, L. 2007. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotechnology,* 25**,** 1298-1306.

LOPES, L., DEWANNIEUX, M., TAKEUCHI, Y. & COLLINS, M. K. 2011. A lentiviral vector pseudotype suitable for vaccine development. *The Journal of Gene Medicine,* 13**,** 181-187.

LU, X., HUMEAU, L., SLEPUSHKIN, V., BINDER, G., YU, Q., SLEPUSHKINA, T., CHEN, Z., MERLING, R., DAVIS, B., CHANG, Y. N. & DROPULIC, B. 2004. Safe two-plasmid production for the first clinical lentivirus vector that achieves >99% transduction in primary cells using a one-step protocol. *Journal of Gene Medicine,* 6**,** 963-73.

LUBAN, J., BOSSOLT, K. L., FRANKE, E. K., KALPANA, G. V. & GOFF, S. P. 1993. Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell,* 73**,** 1067-1078.

LUBAN, J. & GOFF, S. P. 1994. Mutational analysis of cis-acting packaging signals in human immunodeficiency virus type 1 RNA. *Journal of Virology,* 68**,** 3784-93.

MACGREGOR, R. R. 2001. Clinical protocol. A phase 1 open-label clinical trial of the safety and tolerability of single escalating doses of autologous CD4 T cells transduced with VRX496 in HIV-positive subjects. *Human Gene Therapy,* 12**,** 2028-2029.

MADDON, P. J., DALGLEISH, A. G., MCDOUGAL, J. S., CLAPHAM, P. R., WEISS, R. A. & AXEL, R. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell,* 47**,** 333-48.

MAETZIG, T., GALLA, M., BAUM, C. & SCHAMBACH, A. 2011. Gammaretroviral Vectors: Biology, Technology and Application. *Viruses,* 3**,** 677-713.

MALECH, H. L. 2000. Use of serum-free medium with fibronectin fragment enhanced transduction in a system of gas permeable plastic containers to achieve high levels of retrovirus transduction at clinical scale. *Stem Cells,* 18**,** 155-6.

MALECH, H. L., MAPLES, P. B., WHITING-THEOBALD, N., LINTON, G. F., SEKHSARIA, S., VOWELLS, S. J., LI, F., MILLER, J. A., DECARLO, E., HOLLAND, S. M., LEITMAN, S. F.,

CARTER, C. S., BUTZ, R. E., READ, E. J., FLEISHER, T. A., SCHNEIDERMAN, R. D., VAN EPPS, D. E., SPRATT, S. K., MAACK, C. A., ROKOVICH, J. A., COHEN, L. K. & GALLIN, J. I. 1997. Prolonged production of NADPH oxidase-corrected granulocytes after gene therapy of chronic granulomatous disease. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 12133-8.

MALIM, M. H. & BIENIASZ, P. D. 2012. HIV Restriction Factors and Mechanisms of Evasion. *Cold Spring Harbor Perspectives in Medicine,* 2**,** a006940.

MANN, R., MULLIGAN, R. C. & BALTIMORE, D. 1983. CONSTRUCTION OF A RETROVIRUS PACKAGING MUTANT AND ITS USE TO PRODUCE HELPER-FREE DEFECTIVE RETROVIRUS. *Cell,* 33**,** 153-159.

MARIANI, R., CHEN, D., SCHRÖFELBAUER, B., NAVARRO, F., KÖNIG, R., BOLLMAN, B., MÜNK, C., NYMARK-MCMAHON, H. & LANDAU, N. R. 2003. Species-Specific Exclusion of APOBEC3G from HIV-1 Virions by Vif. *Cell,* 114**,** 21-31.

MARTIN-SERRANO, J., YAROVOY, A., PEREZ-CABALLERO, D. & BIENIASZ, P. D. 2003. Divergent retroviral late-budding domains recruit vacuolar protein sorting factors by using alternative adaptor proteins. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 12414-9.

MASUDA, T., KURODA, M. J. & HARADA, S. 1998. Specific and independent recognition of U3 and U5 att sites by human immunodeficiency virus type 1 integrase in vivo. *Journal of Virology,* 72**,** 8396-402.

MAY, C., RIVELLA, S., CALLEGARI, J., HELLER, G., GAENSLER, K. M. L., LUZZATTO, L. & SADELAIN, M. 2000. Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressinglentivirus-encoded human beta-globin. *Nature,* 406**,** 82-86.

MAZARAKIS, N. D., AZZOUZ, M., ROHLL, J. B., ELLARD, F. M., WILKES, F. J., OLSEN, A. L., CARTER, E. E., BARBER, R. D., BABAN, D. F., KINGSMAN, S. M., KINGSMAN, A. J., O'MALLEY, K. & MITROPHANOUS, K. A. 2001. Rabies virus glycoprotein pseudotyping of lentiviral vectors enables retrograde axonal transport and access to the nervous system after peripheral delivery. *Human Molecular Genetics,* 10**,** 2109-21.

MCALLISTER, S. C., HANSEN, S. G., RUHL, R. A., RAGGO, C. M., DEFILIPPIS, V. R., GREENSPAN, D., FRUH, K. & MOSES, A. V. 2004. Kaposi sarcoma-associated herpesvirus (KSHV) induces heme oxygenase-1 expression and activity in KSHV-infected endothelial cells. *Blood,* 103**,** 3465-73.

MERRIL, C. R. & PRATT, M. E. 1986. A silver stain for the rapid quantitative detection of proteins or nucleic acids on membranes or thin layer plates. *Analytical Biochemistry,* 156**,** 96-110.

MERTEN, O.-W., CHARRIER, S., LAROUDIE, N., FAUCHILLE, S., DUGUÉ, C., JENNY, C., AUDIT, M., ZANTA-BOUSSIF, M.-A., CHAUTARD, H., RADRIZZANI, M., VALLANTI, G., NALDINI, L., NOGUIEZ-HELLIN, P. & GALY, A. 2011. Large-Scale Manufacture and Characterization of a Lentiviral Vector Produced for Clinical *Ex Vivo* Gene Therapy Application. *Human Gene Therapy,* 22**,** 343-356.

MILLER, J. C., HOLMES, M. C., WANG, J. B., GUSCHIN, D. Y., LEE, Y. L., RUPNIEWSKI, I., BEAUSEJOUR, C. M., WAITE, A. J., WANG, N. S., KIM, K. A., GREGORY, P. D., PABO, C. O. & REBAR, E. J. 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotechnology,* 25**,** 778-785.

MILLER, M. A., MIETZNER, T. A., CLOYD, M. W., ROBEY, W. G. & MONTELARO, R. C. 1993. Identification of a calmodulin-binding and inhibitory peptide domain in the HIV-1 transmembrane glycoprotein. *AIDS Research and Human Retroviruses,* 9**,** 1057-66.

MITROPHANOUS, K., YOON, S., ROHLL, J., PATIL, D., WILKES, F., KIM, V., KINGSMAN, S., KINGSMAN, A. & MAZARAKIS, N. 1999. Stable gene transfer to the nervous system using a non-primate lentiviral vector. *Gene Therapy,* 6**,** 1808-18.

MIYAUCHI, K., KIM, Y., LATINOVIC, O., MOROZOV, V. & MELIKYAN, G. B. 2009. HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. *Cell,* 137**,** 433-44.

MOERDYK-SCHAUWECKER, M., HWANG, S.-I. & GRDZELISHVILI, V. Z. 2009. Analysis of virion associated host proteins in vesicular stomatitis virus using a proteomics approach. *Virology Journal,* 6**,** 166-166.

MOK, H. P., JAVED, S. & LEVER, A. 2007. Stable gene expression occurs from a minority of integrated HIV-1-based vectors: transcriptional silencing is present in the majority. *Gene Therapy,* 14**,** 741-751.

MOULAND, A. J., MERCIER, J., LUO, M., BERNIER, L., DESGROSEILLERS, L., COHEN, É. A., MOULAND, A. J., MERCIER, J., LUO, M., BERNIER, L., DESGROSEILLERS, L. & COHEN, É. A. 2000. The Double-Stranded RNA-Binding Protein Staufen Is Incorporated in Human Immunodeficiency Virus Type 1: Evidence for a Role in Genomic RNA Encapsidation. *Journal of Virology,* 74**,** 5441-5451.

MUKHERJEE, S. & THRASHER, A. J. 2013. Gene therapy for PIDs: progress, pitfalls and prospects. *Gene,* 525**,** 174-81.

NALDINI, L., BLÖMER, U., GALLAY, P., ORY, D., MULLIGAN, R., GAGE, F. H., VERMA, I. M. & TRONO, D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science,* 272**,** 263-267.

NAUMANN, N., DE RAVIN, S. S., CHOI, U., MOAYERI, M., WHITING-THEOBALD, N., LINTON, G. F., IKEDA, Y. & MALECH, H. L. 2007. Simian immunodeficiency virus lentivector corrects human X-linked chronic granulomatous disease in the NOD/SCID mouse xenograft. *Gene Therapy,* 14**,** 1513-1524.

NEGRE, D., MANGEOT, P. E., DUISIT, G., BLANCHARD, S., VIDALAIN, P. O., LEISSNER, P., WINTER, A. J., RABOURDIN-COMBE, C., MEHTALI, M., MOULLIER, P., DARLIX, J. L. & COSSET, F. L. 2000. Characterization of novel safe lentiviral vectors derived from simian immunodeficiency virus (SIVmac251) that efficiently transduce mature human dendritic cells. *Gene Therapy,* 7**,** 1613-23.

NI, Y. J., SUN, S. S., OPARAOCHA, T., HUMEAU, L., DAVIS, B., COHEN, R., BINDER, G., CHANG, Y. N., SLEPUSHKIN, V. & DROPULIC, B. 2005. Generation of a packaging cell tine for prolonged large-scale production of high-titer HIV-1-based lentiviral vector. *Journal of Gene Medicine,* 7**,** 818-834.

NIIKURA, M., LIU, H. C., DODGSON, J. B. & CHENG, H. H. 2004. A comprehensive screen for chicken proteins that interact with proteins unique to virulent strains of Marek's disease virus. *Poultry Science,* 83**,** 1117-23.

ONAFUWA-NUGA, A. and TELESNITSKY, A. 2009. The Remarkable Frequency of Human Immunodeficiency Virus Type 1 Genetic Recombination. Microbiology and Molecular Biology Reviews, 73, 3, 451–80.

ONO, A. 2010. Relationships between plasma membrane microdomains and HIV-1 assembly. *Biology of the Cell,* 102**,** 335-50.

OTT, D. E. 2002. Potential roles of cellular proteins in HIV-1. *Reviews in Medical Virology,* 12**,** 359-374.

OTT, D. E. 2008. Cellular proteins detected in HIV-1. *Reviews in Medical Virology,* 18**,** 159-175.

OTT, D. E., COREN, L. V., JOHNSON, D. G., SOWDER, R. C., 2ND, ARTHUR, L. O. & HENDERSON, L. E. 1995. Analysis and localization of cyclophilin A found in the virions of human immunodeficiency virus type 1 MN strain. *AIDS Research and Human Retroviruses,* 11**,** 1003-6.

OTT, D. E., COREN, L. V., KANE, B. P., BUSCH, L. K., JOHNSON, D. G., SOWDER, R. C., 2ND, CHERTOVA, E. N., ARTHUR, L. O. & HENDERSON, L. E. 1996. Cytoskeletal proteins inside human immunodeficiency virus type 1 virions. *Journal of Virology,* 70**,** 7734-7743.

OTT, M. G., SCHMIDT, M., SCHWARZWAELDER, K., STEIN, S., SILER, U., KOEHL, U., GLIMM, H., KUHLCKE, K., SCHILZ, A., KUNKEL, H., NAUNDORF, S., BRINKMANN, A., DEICHMANN, A., FISCHER, M., BALL, C., PILZ, I., DUNBAR, C., DU, Y., JENKINS, N. A., COPELAND, N. G., LUTHI, U., HASSAN, M., THRASHER, A. J., HOELZER, D., VON KALLE, C., SEGER, R. & GREZ, M. 2006. Correction of X-linked chronic granulomatous disease by gene therapy,

augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nature Medicine,* 12**,** 401-9.

PADDISON, P. J., CAUDY, A. A., BERNSTEIN, E., HANNON, G. J. & CONKLIN, D. S. 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes & Development,* 16**,** 948-58.

PALFI, S., GURRUCHAGA, J. M., RALPH, G. S., LEPETIT, H., LAVISSE, S., BUTTERY, P. C., WATTS, C., MISKIN, J., KELLEHER, M., DEELEY, S., IWAMURO, H., LEFAUCHEUR, J. P., THIRIEZ, C., FENELON, G., LUCAS, C., BRUGIERES, P., GABRIEL, I., ABHAY, K., DROUOT, X., TANI, N., KAS, A., GHALEH, B., LE CORVOISIER, P., DOLPHIN, P., BREEN, D. P., MASON, S., GUZMAN, N. V., MAZARAKIS, N. D., RADCLIFFE, P. A., HARROP, R., KINGSMAN, S. M., RASCOL, O., NAYLOR, S., BARKER, R. A., HANTRAYE, P., REMY, P., CESARO, P. & MITROPHANOUS, K. A. 2014. Long-term safety and tolerability of ProSavin, a lentiviral vector-based gene therapy for Parkinson's disease: a dose escalation, open-label, phase 1/2 trial. *Lancet*, S0140-6736(13)61939-X.

PANDA, D., DAS, A., DINH, P. X., SUBRAMANIAM, S., NAYAK, D., BARROWS, N. J., PEARSON, J. L., THOMPSON, J., KELLY, D. L., LADUNGA, I. & PATTNAIK, A. K. 2011. RNAi screening reveals requirement for host cell secretory pathway in infection by diverse families of negative-strand RNA viruses. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 19036-41.

PARKER, C.E., MOCANU, V., MOCANU, M., et al. Mass Spectrometry for Post-Translational Modifications. In: Alzate O, editor. *Neuroproteomics*. Boca Raton (FL): CRC Press; 2010. Chapter 6. Available from: http://www.ncbi.nlm.nih.gov/books/NBK56012/

PEAR, W. S., NOLAN, G. P., SCOTT, M. L. & BALTIMORE, D. 1993. Production of high-titer helper-free retroviruses by transient transfection. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 8392-6.

PEARCE-PRATT, R., MALAMUD, D. & M PHILLIPS, D. 1994. Role of the cytoskeleton in cell-to-cell transmission of human immunodeficiency virus. *Journal of Virology,* 68**,** 2898-2905.

PELCHEN-MATTHEWS, A., KRAMER, B. & MARSH, M. 2003. Infectious HIV-1 assembles in late endosomes in primary macrophages. *Journal of Cell Biology,* 162**,** 443-55.

PERUGI, F., MURIAUX, D., RAMIREZ, B. C., CHABANI, S., DECROLY, E., DARLIX, J. L., BLOT, V. & PIQUE, C. 2009. Human Discs Large is a new negative regulator of human immunodeficiency virus-1 infectivity. *Molecular Biology of the Cell,* 20**,** 498-508.

PETROPOULOS, C. Retroviral Taxonomy, Protein Structures, Sequences, and Genetic Maps. In: Coffin JM, Hughes SH, Varmus HE, editors. *Retroviruses*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1997. Available from: http://www.ncbi.nlm.nih.gov/books/NBK19417/

PICHLMAIR, A., DIEBOLD, S. S., GSCHMEISSNER, S., TAKEUCHI, Y., IKEDA, Y., COLLINS, M. K. & REIS E SOUSA, C. 2007. Tubulovesicular Structures within Vesicular Stomatitis Virus G Protein-Pseudotyped Lentiviral Vector Preparations Carry DNA and Stimulate Antiviral Responses via Toll-Like Receptor 9. *Journal of Virology,* 81**,** 539-547.

POESCHLA, E., GILBERT, J., LI, X., HUANG, S., HO, A. & WONG-STAAL, F. 1998. Identification of a human immunodeficiency virus type 2 (HIV-2) encapsidation determinant and transduction of nondividing human cells by HIV-2-based lentivirus vectors. *Journal of Virology,* 72**,** 6527-36.

RAIBORG, C. & STENMARK, H. 2009. The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature,* 458**,** 445-452.

RAMAGLI, L. S. & RODRIGUEZ, L. V. 1985. Quantitation of microgram amounts of protein in two-dimensional polyacrylamide gel electrophoresis sample buffer. *ELECTROPHORESIS,* 6**,** 559-563.

REED, S. E., STALEY, E. M., MAYGINNES, J. P., PINTEL, D. J. & TULLIS, G. E. 2006. Transfection of mammalian cells using linear polyethylenimine is a simple and effective means of

producing recombinant adeno-associated virus vectors. *Journal of Virological Methods,* 138**,** 85-98.

REEVES, R. H. & O'BRIEN, S. J. 1984. Molecular genetic characterization of the RD-114 gene family of endogenous feline retroviral sequences. *Journal of Virology,* 52**,** 164-71.

RELANDER, T., JOHANSSON, M., OLSSON, K., IKEDA, Y., TAKEUCHI, Y., COLLINS, M. & RICHTER, J. 2005. Gene Transfer to Repopulating Human CD34+ Cells Using Amphotropic-, GALV-, or RD114-Pseudotyped HIV-1-Based Vectors from Stable Producer Cells. *Molecular Therapy: The Journal of the American Society of Gene Therapy,* 11**,** 452-459.

RODRIGUES, T., ALVES, A., LOPES, A., CARRONDO, M. J. T., ALVES, P. M. & CRUZ, P. E. 2008. Removal of envelope protein-free retroviral vectors by anion-exchange chromatography to improve product quality. *Journal of Separation Science,* 31**,** 3509-3518.

RODRIGUEZ, M., SARRY, E., BEJANARIU, A., CASABAN, L., ABDALLA, S., SABBAH-PETROVER, E. & BAUCHE, C. 2012. Design and development of a new lentiviral based anti-HIV therapeutic vaccine. *Retrovirology,* 9.

ROWE, H. M., LOPES, L., BROWN, N., EFKLIDOU, S., SMALLIE, T., KARRAR, S., KAYE, P. M. & COLLINS, M. K. 2009. Expression of vFLIP in a Lentiviral Vaccine Vector Activates NF-kappa B, Matures Dendritic Cells, and Increases CD8(+) T-Cell Responses. *Journal of Virology,* 83**,** 1555-1562.

ROWE, H. M., LOPES, L., IKEDA, Y., BAILEY, R., BARDE, I., ZENKE, M., CHAIN, B. M. & COLLINS, M. K. 2006. Immunization with a lentiviral vector stimulates both CD4 and CD8 T cell responses to an ovalbumin transgene. *Molecular Therapy,* 13**,** 310-9.

RUIZ-MAY, E., THANNHAUSER, T. W., ZHANG, S. & ROSE, J. K. 2012. Analytical technologies for identification and characterization of the plant N-glycoproteome. *Frontiers in Plant Science,* 3**,** 150.

SAAD, J. S., MILLER, J., TAI, J., KIM, A., GHANAM, R. H., SUMMERS, M. F., SUMMERS@HHMI.UMBC.EDU, SAAD, J. S., MILLER, J., TAI, J., KIM, A., GHANAM, R. H. & SUMMERS, M. F. 2006. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 11364-11369.

SAKUMA, T., BARRY, M. A. & IKEDA, Y. 2012. Lentiviral vectors: basic to translational. *Biochemistry Journal,* 443**,** 603-18.

SANDRIN, V., BOSON, B., SALMON, P., GAY, W., NÈGRE, D., GRAND, R. L., TRONO, D. & COSSET, F.-L. C. 2002. Lentiviral vectors pseudotyped with a modified RD114 envelope glycoprotein show increased stability in sera and augmented transduction of primary lymphocytes and CD34+ cells derived from human and nonhuman primates. *Blood,* 100**,** 823-832.

SANTOS, S., OBUKHOV, Y., NEKHAI, S., BUKRINSKY, M. & IORDANSKIY, S. 2012. Virus-producing cells determine the host protein profiles of HIV-1 virion cores. *Retrovirology,* 9**,** 65.

SAPHIRE, A. C., GALLAY, P. A. & BARK, S. J. 2006a. Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *Journal of Proteome Research,* 5**,** 530-538.

SAPHIRE, A. C., GALLAY, P. A. & BARK, S. J. 2006b. Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *J Proteome Res,* 5**,** 530-8.

SASTRY, L., XU, Y., COOPER, R., POLLOK, K. & CORNETTA, K. 2004. Evaluation of plasmid DNA removal from lentiviral vectors by benzonase treatment. *Human Gene Therapy,* 15**,** 221-6.

SATTENTAU, Q. 2008. Avoiding the void: cell-to-cell spread of human viruses. *Nature Review Microbiolgy,* 6**,** 815-26.

SCHALLER, T., OCWIEJA, K., RASAIYAAH, J., PRICE, A., BRADY, T., ROTH, S., HUÉ, S., FLETCHER, A., LEE, K., KEWALRAMANI, V., NOURSADEGHI, M., JENNER, R., JAMES, L., BUSHMAN,

F. & TOWERS, G. 2011a. HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathogens,* 7**,** e1002439.

SCHALLER, T., OCWIEJA, K., RASAIYAAH, J., PRICE, A., BRADY, T., ROTH, S., HUÉ, S., FLETCHER, A., LEE, K., KEWALRAMANI, V., NOURSADEGHI, M., JENNER, R., JAMES, L., BUSHMAN, F. & TOWERS, G. 2011b. HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathogens,* 7**,** e1002439.

SCHAMBACH, A., MUELLER, D., GALLA, M., VERSTEGEN, M.M.A., WAGEMAKER G., LOEW, R., BAUM, C. AND BOHNE, J., 2006. Overcoming promoter competition in packaging cells improves production of self-inactivating retroviral vectors. Gene Therapy, 13, 1524–1533.

SCHAMBACH, A., GALLA, M., MAETZIG, T., LOEW, R. & BAUM, C. 2007. Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors. *Molecular Therapy,* 15**,** 1167-73.

SCHLEGEL, R., TRALKA, T. S., WILLINGHAM, M. C. & PASTAN, I. 1983. Inhibition of VSV binding and infectivity by phosphatidylserine: is phosphatidylserine a VSV-binding site? *Cell,* 32**,** 639-46.

SCHRÖDER, A. R. W., SHINN, P., CHEN, H., BERRY, C., ECKER, J. R. & BUSHMAN, F. 2002. HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. *Cell,* 110**,** 521-529.

SCHUCHT, R., COROADINHA, A. S., ZANTA-BOUSSIF, M. A., VERHOEYEN, E., CARRONDO, M. J. T., HAUSER, H. & WIRTH, D. 2006. A New Generation of Retroviral Producer Cells: Predictable and Stable Virus Production by Flp-Mediated Site-Specific Integration of Retroviral Vectors. *Molecular Therapy: The Journal of the American Society of Gene Therapy,* 14**,** 285-292.

SCHWARTZ, S., FELBER, B. K. & PAVLAKIS, G. N. 1992. Distinct RNA sequences in the gag region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. *Journal of Virology,* 66**,** 150-9.

SCHWEIZER, M. & MERTEN, O. W. 2010. Large-Scale Production Means for the Manufacturing of Lentiviral Vectors. *Current Gene Therapy,* 10**,** 474-486.

SEGURA, M. D. L. M., GARNIER, A. & KAMEN, A. 2006a. Purification and characterization of retrovirus vector particles by rate zonal ultracentrifugation. *Journal of Virological Methods,* 133**,** 82-91.

SEGURA, M. D. L. M., KAMEN, A. & GARNIER, A. 2006b. Downstream processing of oncoretroviral and lentiviral gene therapy vectors. *Biotechnology Advances,* 24**,** 321-337.

SEGURA, M. D. L. M., KAMEN, A., TRUDEL, P. & GARNIER, A. 2005. A novel purification strategy for retrovirus gene therapy vectors using heparin affinity chromatography. *Biotechnology and Bioengineering,* 90**,** 391-404.

SEGURA, M. M., GARNIER, A., DI FALCO, M. R., WHISSELL, G., MENESES-ACOSTA, A., ARCAND, N. & KAMEN, A. 2008. Identification of Host Proteins Associated with Retroviral Vector Particles by Proteomic Analysis of Highly Purified Vector Preparations. *Journal of Virology,* 82**,** 1107-1117.

SEGURA, M. M., GARNIER, A., DUROCHER, Y., COELHO, H. & KAMEN, A. 2007. Production of lentiviral vectors by large-scale transient transfection of suspension cultures and affinity chromatography purification. *Biotechnology and Bioengineering,* 98**,** 789-799.

SEGURA, M. M., MANGION, M., GAILLET, B. & GARNIER, A. 2013. New developments in lentiviral vector design, production and purification. *Expert Opinion on Biological Therapy,* 13**,** 987-1011.

SEKIYA, F., BAE, Y. S., JHON, D. Y., HWANG, S. C. & RHEE, S. G. 1999. AHNAK, a protein that binds and activates phospholipase C-gamma1 in the presence. *Journal of Biological Chemistry,* 274**,** 13900-7.

SHAW, M. L., STONE, K. L., COLANGELO, C. M., GULCICEK, E. E. & PALESE, P. 2008. Cellular proteins in influenza virus particles. *PLoS Pathogens,* 4**,** e1000085.

SHEEHY, A. M., GADDIS, N. C., CHOI, J. D. & MALIM, M. H. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature,* 418**,** 646-50.

SHIMOJIMA, M., MIYAZAWA, T., IKEDA, Y., MCMONAGLE, E. L., HAINING, H., AKASHI, H., TAKEUCHI, Y., HOSIE, M. J. & WILLETT, B. J. 2004. Use of CD134 as a primary receptor by the feline immunodeficiency virus. *Science,* 303**,** 1192-5.

SHTIVELMAN, E., COHEN, F. E. & BISHOP, J. M. 1992. A human gene (AHNAK) encoding an unusually large protein with a 1.2-microns polyionic rod structure. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 5472-6.

SIMONS, K. & IKONEN, E. 1997. Functional rafts in cell membranes. *Nature,* 387**,** 569-72.

SLEPUSHKIN, V., CHANG, N., COHEN, R. & GAN, Y. 2003. Large scale Purification of a lentiviral vector by SEC or Mustang Q Ion Exchange Capsule. *Bioprocessing Journal,* 2**,** 89-95.

SOKOLSKAJA, E. & LUBAN, J. 2006. Cyclophilin, TRIM5, and innate immunity to HIV-1. *Current Opinion in Microbiology,* 9**,** 404-8.

SORIN, M., YUNG, E., WU, X. & KALPANA, G. V. 2006. HIV-1 replication in cell lines harboring INI1/hSNF5 mutations. *Retrovirology,* 3.

SOWINSKI, S., JOLLY, C., BERNINGHAUSEN, O., PURBHOO, M. A., CHAUVEAU, A., KOHLER, K., ODDOS, S., EISSMANN, P., BRODSKY, F. M., HOPKINS, C., ONFELT, B., SATTENTAU, Q. & DAVIS, D. M. 2008. Membrane nanotubes physically connect T cells over long distances presenting a novel route for HIV-1 transmission. *Nature Cell Biology,* 10**,** 211-219.

STEIN, S., OTT, M. G., SCHULTZE-STRASSER, S., JAUCH, A., BURWINKEL, B., KINNER, A., SCHMIDT, M., KRÄMER, A., SCHWÄBLE, J., GLIMM, H., KOEHL, U., PREISS, C., BALL, C., MARTIN, H., GÖHRING, G., SCHWARZWAELDER, K., HOFMANN, W.-K., KARAKAYA, K., TCHATCHOU, S., YANG, R., REINECKE, P., KÜHLCKE, K., SCHLEGELBERGER, B., THRASHER, A. J., HOELZER, D., SEGER, R., KALLE, C. V. & GREZ, M. 2010. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nature Medicine,* 16**,** 198-204.

STEWART, H. J., FONG-WONG, L., STRICKLAND, I., CHIPCHASE, D., KELLEHER, M., STEVENSON, L., THOREE, V., MCCARTHY, J., RALPH, G. S., MITROPHANOUS, K. & RADCLIFFE, P. A. 2011. A Stable Producer Cell Line for the Manufacture of a Lentiviral Vector for Gene Therapy of Parkinson's Disease. *Human Gene Therapy,* 22**,** 357-369.

STEWART, H. J., LEROUX-CARLUCCI, M. A., SION, C. J. M., MITROPHANOUS, K. A. & RADCLIFFE, P. A. 2009. Development of inducible EIAV-based lentiviral vector packaging and producer cell lines. *Gene Therapy,* 16**,** 805-814.

STRACK, B., CALISTRI, A., CRAIG, S., POPOVA, E. & GÖTTLINGER, H. G. 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell,* 114**,** 689-99.

STRANG, B. L., IKEDA, Y., COSSET, F. L., COLLINS, M. K. L. & TAKEUCHI, Y. 2004. Characterization of HIV-1 vectors with gammaretrovirus envelope glycoproteins produced from stable packaging cells. *Gene Therapy,* 11**,** 591-598.

STREMLAU, M., OWENS, C. M., PERRON, M. J., KIESSLING, M., AUTISSIER, P. & SODROSKI, J. 2004. The cytoplasmic body component TRIM5 alpha restricts HIV-1 infection in Old World monkeys. *Nature,* 427**,** 848-853.

SUZUKI, Y. & CRAIGIE, R. 2007. The road to chromatin - nuclear entry of retroviruses. *Nature Reviews Microbiology,* 5**,** 187-196.

SWANSON, C. M. & MALIM, M. H. 2008. SnapShot: HIV-1 proteins. *Cell,* 133**,** 742, 742.e1.

SWITZER C ROBERT, M. R. C., SHIFRIN SIDNEY 1979. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. 98**,** 231–237.

TAILOR, C. S., NOURI, A., ZHAO, Y., TAKEUCHI, Y. & KABAT, D. 1999. A sodium-dependent neutral-amino-acid transporter mediates infections of feline and baboon endogenous retroviruses and simian type D retroviruses. *Journal of Virology,* 73**,** 4470-4.

TAKAHASHI, N., HAYANO, T. & SUZUKI, M. 1989. Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin. *Nature,* 337**,** 473-5.

TANG, C., LOELIGER, E., LUNCSFORD, P., KINDE, I., BECKETT, D. & SUMMERS, M. F. 2004. Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 517-22.

TEBAS, P., STEIN, D., BINDER-SCHOLL, G., MUKHERJEE, R., BRADY, T., REBELLO, T., HUMEAU, L., KALOS, M., PAPASAVVAS, E., MONTANER, L. J., SCHULLERY, D., SHAHEEN, F., BRENNAN, A. L., ZHENG, Z., COTTE, J., SLEPUSHKIN, V., VELOSO, E., MACKLEY, A., HWANG, W.-T., ABERRA, F., ZHAN, J., BOYER, J., COLLMAN, R. G., BUSHMAN, F. D., LEVINE, B. L. & JUNE, C. H. 2013. Antiviral effects of autologous CD4 T cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to HIV. *Blood,* 121**,** 1524-1533.

THE_JOURNAL_OF_GENE_MEDICINE. 2014. *Gene Therapy Clinical Trials Worldwide* [Online]. Available: http://www.abedia.com/wiley/vectors.php.

THRASHER, A. J., HACEIN-BEY-ABINA, S., GASPAR, H. B., BLANCHE, S., DAVIES, E. G., PARSLEY, K., GILMOUR, K., KING, D., HOWE, S., SINCLAIR, J., HUE, C., CARLIER, F., VON KALLE, C., DE SAINT BASILE, G., LE DEIST, F., FISCHER, A. & CAVAZZANA-CALVO, M. 2005. Failure of SCID-X1 gene therapy in older patients. *Blood,* 105**,** 4255-7.

THROM, R. E., OUMA, A. A., ZHOU, S., CHANDRASEKARAN, A., LOCKEY, T., GREENE, M., DE RAVIN, S. S., MOAYERI, M., MALECH, H. L., SORRENTINO, B. P. & GRAY, J. T. 2009. Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection. *Blood,* 113**,** 5104-10.

TOLEDO, J. R., PRIETO, Y., ORAMAS, N. & SANCHEZ, O. 2009. Polyethylenimine-based transfection method as a simple and effective way to produce recombinant lentiviral vectors. *Applied Biochemistry and Biotechnology,* 157**,** 538-44.

TOWERS, G. J. 2007. The control of viral infection by tripartite motif proteins and cyclophilin A. *Retrovirology,* 4**,** 40-40.

TRANSFIGURACION, J., JAALOUK, D. E., GHANI, K., GALIPEAU, J. & KAMEN, A. 2003. Size-exclusion chromatography purification of high-titer vesicular stomatitis virus G glycoprotein-pseudotyped retrovectors for cell and gene therapy applications. *Human Gene Therapy,* 14**,** 1139-1153.

UCHIDA, N., WASHINGTON, K. N., LAP, C. J., HSIEH, M. M. & TISDALE, J. F. 2011. Chicken HS4 Insulators Have Minimal Barrier Function Among Progeny of Human Hematopoietic Cells Transduced With an HIV1-based Lentiviral Vector. *Molecular Therapy,* 19**,** 133-139.

USAMI, Y., POPOV, S. & GOTTLINGER, H. G. 2007. Potent rescue of human immunodeficiency virus type 1 late domain mutants by ALIX/AIP1 depends on its CHMP4 binding site. *Journal of Virology,* 81**,** 6614-22.

VAN LINT, C., BOUCHAT, S. & MARCELLO, A. 2013. HIV-1 transcription and latency: an update. *Retrovirology,* 10**,** 38.

VARNUM, S. M., STREBLOW, D. N., MONROE, M. E., SMITH, P., AUBERRY, K. J., PASA-TOLIC, L., WANG, D., CAMP, D. G., RODLAND, K., WILEY, S., BRITT, W., SHENK, T., SMITH, R. D. & NELSON, J. A. 2004. Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome. *Journal of Virology*, 78, 10960-6.

VERMA, I. M. 2013. Gene Therapy That Works. *Science*, 341, 853-5

VERONESE, F. D., COPELAND, T. D., OROSZLAN, S., GALLO, R. C. & SARNGADHARAN, M. G. 1988. Biochemical and immunological analysis of human immunodeficiency virus gag gene products p17 and p24. *Journal of Virology*, 62, 795-801

VERPLANK, L., BOUAMR, F., LAGRASSA, T. J., AGRESTA, B., KIKONYOGO, A., LEIS, J. & CARTER, C. A. 2001. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 7724-9.

VOTTELER, J. & SUNDQUIST, W. I. 2013. Virus Budding and the ESCRT Pathway. *Cell Host and Microbe,* 14**,** 232-41.

WARREN, K., WEI, T., LI, D., QIN, F., WARRILOW, D., LIN, M. H., SIVAKUMARAN, H., APOLLONI, A., ABBOTT, C. M., JONES, A., ANDERSON, J. L. & HARRICH, D. 2012. Eukaryotic elongation factor 1 complex subunits are critical HIV-1 reverse transcription cofactors. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 9587-92.

WASHBURN, M. P., WOLTERS, D. & YATES, J. R. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology,* 19**,** 242-247.

WEISS, E. R. & GÖTTLINGER, H. 2011. The role of cellular factors in promoting HIV budding. *Journal of Molecular Biology,* 410**,** 525-533.

WELTE, M. A. 2004. Bidirectional transport along microtubules. *Current Biology,* 14**,** R525-R537.

WHEELER, J. X., JONES, C., THORPE, R. & ZHAO, Y. 2007. Proteomics analysis of cellular components in lentiviral vector production using Gel-LC-MS/MS. *PROTEOMICS – Clinical Applications,* 1**,** 224-230.

WILK, T., GOWEN, B. & FULLER, S. D. 1999. Actin associates with the nucleocapsid domain of the human immunodeficiency virus Gag polyprotein. *Journal of Virology,* 73**,** 1931-40.

WOJCECHOWSKYJ, J. A., DIDIGU, C. A., LEE, J. Y., PARRISH, N. F., SINHA, R., HAHN, B. H., BUSHMAN, F. D., JENSEN, S. T., SEEHOLZER, S. H. & DOMS, R. W. 2013. Quantitative phosphoproteomics reveals extensive cellular reprogramming during HIV-1 entry. *Cell Host Microbe,* 13**,** 613-23.

WOODWARD, C. L., PRAKOBWANAKIT, S., MOSESSIAN, S. & CHOW, S. A. 2009. Integrase Interacts with Nucleoporin NUP153 To Mediate the Nuclear Import of Human Immunodeficiency Virus Type 1. *Journal of Virology,* 83**,** 6522-33.

WORLD_HEALTH_ORGANIZATION 2014. WHO Antiretroviral Therapy Guidlines 2013 - Summary of new recommendations. *WHO*.

WU, Y. 2004. HIV-1 gene expression: lessons from provirus and non-integrated DNA. *Retrovirology,* 1**,** 13.

XU, K. L., MA, H., MCCOWN, T. J., VERMA, I. M. & KAFRI, T. 2001. Generation of a stable cell line producing high-titer self-inactivating lentiviral vectors. *Molecular Therapy,* 3**,** 97-104.

YAMADA, K., MCCARTY, D. M., MADDEN, V. J. & WALSH, C. E. 2003. Lentivirus vector purification using anion exchange HPLC leads to improved gene transfer. *BioTechniques,* 34**,** 1074-1078, 1080-1074-1078, 1080.

YANEZ-MUNOZ, R. J., BALAGGAN, K. S., MACNEIL, A., HOWE, S. J., SCHMIDT, M., SMITH, A. J., BUCH, P., MACLAREN, R. E., ANDERSON, P. N., BARKER, S. E., DURAN, Y., BARTHOLOMAE, C., VON KALLE, C., HECKENLIVELY, J. R., KINNON, C., ALI, R. R. & THRASHER, A. J. 2006. Effective gene therapy with nonintegrating lentiviral vectors. *Nature Medicine,* 12**,** 348-353.

YANG, H. G., HU, B. L., XIAO, L. & WANG, P. 2011. Dendritic cell-directed lentivector vaccine induces antigen-specific immune responses against murine melanoma. *Cancer Gene Therapy, 18, 370-80*

YI, R., QIN, Y., MACARA, I. G. & CULLEN, B. R. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes and Development*, 17, 3011-16

YU, S. F., VON RUDEN, T., KANTOFF, P. W., GARBER, C., SEIBERG, M., RUTHER, U., ANDERSON, W. F., WAGNER, E. F. & GILBOA, E. 1986. Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America,* 83**,** 3194-8.

ZAISS, A. K., SON, S. & CHANG, L. J. 2002. RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. *Journal of Virology,* 76**,** 7209-19.

ZAITSEVA, L., MYERS, R. & FASSATI, A. 2006. tRNAs promote nuclear import of HIV-1 intracellular reverse transcription complexes. *PLoS Biology,* 4**,** e332.

ZENNOU, V., PETIT, C., GUETARD, D., NERHBASS, U., MONTAGNIER, L. & CHARNEAU, P. 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell,* 101**,** 173-85.

ZHANG, F., FROST, A. R., BLUNDELL, M. P., BALES, O., ANTONIOU, M. N. & THRASHER, A. J. 2010. A Ubiquitous Chromatin Opening Element (UCOE) Confers Resistance to DNA Methylation-mediated Silencing of Lentiviral Vectors. *Molecular Therapy,* 18**,** 1640-1649.

ZHANG, F., THORNHILL, S. I., HOWE, S. J., ULAGANATHAN, M., SCHARNBACH, A., SINCLAIR, J., KINNON, C., GASPAR, H. B., ANTONIOU, M. & THRASHER, A. J. 2007. Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells. *Blood,* 110**,** 1448-1457.

ZHOU, H., XU, M., HUANG, Q., GATES, A. T., ZHANG, X. D., CASTLE, J. C., STEC, E., FERRER, M., STRULOVICI, B., HAZUDA, D. J. & ESPESETH, A. S. 2008. Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe,* 4**,** 495-504.

ZHU, F. X., CHONG, J. M., WU, L. & YUAN, Y. 2005. Virion proteins of Kaposi's sarcoma-associated herpesvirus. *Journal of Virology,* 79**,** 800-11.

ZHU, P., CHERTOVA, E., BESS, J., LIFSON, J. D., ARTHUR, L. O., LIU, J., TAYLOR, K. A. & ROUX, K. H. 2003. Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 15812-7.

ZUFFEREY, R., DONELLO, J. E., TRONO, D. & HOPE, T. J. 1999. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *Journal of Virology,* 73**,** 2886-2892.

ZUFFEREY, R., NAGY, D., J. MANDEL, R., NALDINI, L. & TRONO, D. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nature Biotechnology,* 15**,** 871-875.

6